

# Clustering de séquences d'activités pour l'étude de procédures neurochirurgicales

Germain Forestier\*, Florent Lalys\*\*, Laurent Riffaud\*\*, Brivael Trelhu\*\*, Pierre Jannin\*\*

\*ENSISA/MIPS/Université de Haute Alsace

\*\*INRIA/INSERM/CNRS/Université de Rennes I/IRISA

**Résumé.** L'utilisation de modèles de procédure chirurgicale (Surgical Process Model, SPM) a récemment émergé dans le domaine de la conception d'outils d'intervention chirurgicale assistée par ordinateur. Ces modèles, qui sont utilisés pour analyser et évaluer les interventions, représentent des procédures chirurgicales (Surgical Process, SP) qui sont formalisées comme des structures symboliques décrivant une chirurgie à un niveau de granularité donné. Un enjeu important réside dans la définition de métriques permettant la comparaison et l'évaluation de ces procédures. Ainsi, les relations entre ces métriques et des données pré-opératoires permettent de classer les chirurgies pour mettre en lumière des informations sur la procédure elle-même, mais également sur le comportement du chirurgien. Dans ce papier, nous étudions la classification automatique d'un ensemble de procédures chirurgicales en utilisant l'algorithme Dynamic Time Warping (DTW) pour calculer une mesure de similarité entre procédures chirurgicales. L'utilisation de DTW permet de se concentrer sur les différents types d'activité effectués pendant la procédure, ainsi que sur leur séquençement tout en réduisant les différences temporelles. Des expériences ont été menées sur 24 procédures chirurgicales d'hernie discale lombaire dans le but de discriminer le niveau d'expertise des chirurgiens à partir d'une classification connue. A l'aide d'un algorithme de clustering hiérarchique utilisant DTW nous avons retrouvé deux groupes de chirurgiens présentant des niveaux d'expertise différents (junior et senior).

## 1 Introduction

Dans le domaine de l'ingénierie médicale, l'analyse et la modélisation des procédures chirurgicales sont des sujets qui ont récemment émergé. Les procédures chirurgicales peuvent être décomposées en quatre niveaux de granularité (Mackenzie et al., 2001) : les phases, étapes, tâches et mouvement. Une intervention chirurgicale peut alors être décrite en utilisant un niveau de granularité prédéfini afin de créer une *Surgical Process*, appelé SP dans la suite de cet article. Ces SPs peuvent être vues comme la séquence des activités effectuées par le chirurgien lors d'une intervention. De fait, un enjeu important réside dans la définition de méthodes permettant de comparer et de grouper automatiquement des SPs similaires afin d'identifier des motifs qui peuvent par la suite être mis en relation avec des données pré-opératoires afin d'analyser de manière précise un type de chirurgie.

Le principal problème d'une telle analyse est la définition de mesures de similarité entre SPs qui permettent de mettre en lumière de manière objective et quantitative les différences entre des procédures chirurgicales. En effet, les SPs d'un même type d'intervention peuvent avoir une variabilité importante, qui peut être due à différents paramètres tel que la technique d'opération, la difficulté intrinsèque de la chirurgie ou encore l'expertise du chirurgien. Récemment, Riffaud et al. (2010) ont proposé la définition de métriques de similarité calculées à l'aide d'outils statistiques pour comparer des groupes de chirurgiens junior et senior, c'est-à-dire avec peu ou beaucoup d'expérience. Faisant suite à ces travaux, nous proposons dans cet article de définir une nouvelle approche utilisant l'algorithme Dynamic Time Warping (DTW) (Hiroaki et Chiba, 1978) afin d'évaluer la similarité entre SPs. L'utilisation de DTW nous permet de nous concentrer sur le séquençement des différentes activités qui composent une SPs. En effet, DTW réduit l'importance des différences temporelles et se concentre sur le nombre d'activités et leur organisation dans la chronologie de la chirurgie. A l'aide de DTW, nous avons étudié la classification automatique d'un ensemble de SPs à travers l'utilisation d'un algorithme de clustering hiérarchique ascendant. Ainsi, nous présentons des résultats pour le clustering de 24 SPs correspondant à des chirurgies d'hernie lombaire discale pratiquées pour moitié par des chirurgiens juniors et pour moitié par des chirurgiens seniors. Les résultats montrent que notre approche a réussi de manière automatique à retrouver ces deux groupes de chirurgiens.

Dans la Section 2, le formalisme et l'utilisation de DTW pour la comparaison de procédures sont présentés. La Section 3 présente les résultats des expériences, et la Section 4 conclut l'article.

## 2 Méthodes

### 2.1 Séquence d'activités composant le processus chirurgical

Une procédure chirurgicale (Surgical Process, SP) peut être vue comme une séquence représentant les *activités*  $\mathbf{ac}_i$  effectuées par le chirurgien pendant la chirurgie. L'ensemble de ces activités étant noté  $\mathcal{AC}$  avec  $\mathbf{ac}_i \in \mathcal{AC}$  ( $\mathbf{ac}_i$  étant la  $i^{\text{ème}}$  activité). Chaque activité dans une SP contient plusieurs informations, représentées par un triplet :

$$\mathbf{ac}_i = \langle \mathbf{a}; \mathbf{s}; \mathbf{i} \rangle \quad \mathbf{a} \in \mathcal{A}, \mathbf{s} \in \mathcal{S}, \mathbf{i} \in \mathcal{I}^{m_i} \quad (1)$$

avec  $\mathcal{A}$  l'ensemble des actions possibles (e.g. {cut, remove, ...}),  $\mathcal{S}$  l'ensemble des structures anatomiques (e.g. {skin, dura matter ...}),  $\mathcal{I}$  l'ensemble des instruments possibles (e.g. {scalpel, scissors, ...}) et  $m_i$  le nombre d'instruments utilisés pour réaliser l'activité  $\mathbf{ac}_i$ . Un exemple d'activité peut être :  $\langle \text{cut}, \text{skin}, \text{scalpel} \rangle$ . Ainsi, le domaine de définition d'une activité est donné par :  $\mathcal{A} \times \mathcal{S} \times \mathcal{I}^{m_i}$ . Ces ensembles sont généralement spécifiques à un type de chirurgie, et une ontologie peut être utilisée pour décrire le vocabulaire pour un type défini (Neumuth et al., 2010). En plus des informations sur l'action ( $\mathbf{a}$ ), la structure anatomique ( $\mathbf{s}$ ) et l'instrument utilisé ( $\mathbf{i}$ ), chaque activité a un début ( $start(\mathbf{ac}_i)$ ) et une fin ( $stop(\mathbf{ac}_i)$ ) dans la chronologie de la chirurgie ( $start(.) \rightarrow \mathbb{R}, stop(.) \rightarrow \mathbb{R}$ ). Notons que  $start(\mathbf{ac}_i) < stop(\mathbf{ac}_i)$  induit un ordre partiel sur les activités.

Une procédure chirurgicale ( $\mathbf{sp}_k$ ) peut ainsi être vue comme une séquence d'activités pratiquées pendant une chirurgie. Chacune des activités appartient à l'ensemble des différentes activités présentes pendant la chirurgie ( $\mathcal{AC}_k$ ) :

$$\mathbf{sp}_k = \langle \mathbf{ac}_1^{(k)}, \mathbf{ac}_2^{(k)}, \dots, \mathbf{ac}_{n_k}^{(k)} \rangle \mid \mathbf{ac}_i^{(k)} \in \mathcal{AC}_k \quad (2)$$

## 2.2 Comparaison de procédures à l'aide de Dynamic Time Warping (DTW)

Un enjeu important réside dans la définition de métriques permettant d'évaluer la similarité entre SPs. En effet, définir une mesure de similarité est souvent la première étape pour identifier des motifs parmi un ensemble d'objets. Comme une SP peut être vue comme une séquence d'activités, nous proposons d'utiliser l'algorithme Dynamic Time Warping (DTW) (Hiroaki et Chiba, 1978) afin de les comparer. DTW est basé sur la distance de Levenshtein (ou distance d'édition), et a été initialement utilisé dans des applications de reconnaissance de voix. Il permet de trouver l'alignement optimal entre deux séquences, et de capturer la similarité lors de l'alignement de ces deux séquences. Afin de pouvoir utiliser DTW pour comparer deux séquences, une mesure de similarité doit être définie entre les différents éléments composant la séquence. Dans notre cas, cela implique la définition d'une distance entre deux activités. Ainsi, nous avons défini une fonction binaire renvoyant 0 si les trois composantes (Eq. 1) des deux activités sont identiques, et 1 sinon :

$$d(\mathbf{ac}_i, \mathbf{ac}_j) = \begin{cases} 0 & \text{if } \mathbf{ac}_i(\mathbf{a}) \stackrel{*}{=} \mathbf{ac}_j(\mathbf{a}) \text{ et} \\ & \mathbf{ac}_i(\mathbf{s}) \stackrel{*}{=} \mathbf{ac}_j(\mathbf{s}) \text{ et} \\ & \mathbf{ac}_i(\mathbf{i}) \stackrel{*}{=} \mathbf{ac}_j(\mathbf{i}) \\ 1 & \text{sinon} \end{cases} \quad (3)$$

avec  $\stackrel{*}{=}$  un opérateur booléen qui réalise la comparaison entre l'action, la structure anatomique et l'instrument utilisé (e.g.  $d(\langle \text{cut, skin, scalpel} \rangle, \langle \text{cut, skin, scalpel} \rangle) = 0$ ).

Pour comparer deux SPs en utilisant DTW, les séquences d'activités sont étirées en considérant le début et la fin de chaque activité. Cette étape est nécessaire pour avoir les SPs sur la même échelle temporelle et pouvoir ainsi comparer les activités effectuées dans chaque SP à un instant donné  $t$  pendant la chirurgie. L'activité pratiquée à l'instant  $t$  est définie comme  $\mathbf{ac}_i(t)$  ssi  $t \in [\text{start}(\mathbf{ac}_i); \text{stop}(\mathbf{ac}_i)]$ . Notons qu'il n'est pas nécessaire que les deux SPs aient la même durée, la seule hypothèse étant que la première activité ait commencé à l'instant  $t = 0$  dans les deux SPs.

Considérons deux SPs  $\mathbf{sp}_k = \langle \mathbf{ac}_1^{(k)}, \mathbf{ac}_2^{(k)}, \dots, \mathbf{ac}_{n_k}^{(k)} \rangle$  et  $\mathbf{sp}_l = \langle \mathbf{ac}_1^{(l)}, \mathbf{ac}_2^{(l)}, \dots, \mathbf{ac}_{n_l}^{(l)} \rangle$ , le coup de l'alignement optimal peut être calculé récursivement par :

$$d(\mathbf{sp}_k(t), \mathbf{sp}_l(t)) = d(\mathbf{ac}_i^{(k)}(t), \mathbf{ac}_j^{(l)}(t)) + \min \begin{cases} d(\mathbf{sp}_k(t-1), \mathbf{sp}_l(t-1)) \\ d(\mathbf{sp}_k(t), \mathbf{sp}_l(t-1)) \\ d(\mathbf{sp}_k(t-1), \mathbf{sp}_l(t)) \end{cases} \quad (4)$$

où  $\mathbf{sp}_k(t)$  est la sous-séquence  $\langle \mathbf{ac}_1^{(k)}, \dots, \mathbf{ac}_i^{(k)}(t) \rangle$ .

DTW a déjà été utilisé avec succès comme mesure de similarité pour effectuer du clustering (Petitjean et al., 2011). De manière similaire, nous proposons d'utiliser la mesure de similarité fournie par DTW pour identifier automatiquement des clusters de chirurgies similaires. Nous

## Clustering de séquences d'activités

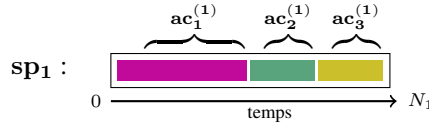


FIG. 1 – Exemple de SP avec trois activités.

avons ainsi utilisé un algorithme de clustering hiérarchique ascendant avec la méthode du lien-moyen (*average-link*) (Anderberg, 1973), qui consiste à évaluer la similarité entre deux clusters par la similarité moyenne entre les couples d'objets des deux clusters.

### 2.3 Visualisation des procédures chirurgicales

Il est en général utile de posséder une représentation visuelle des données pour pouvoir les explorer. Cependant, les structures de données complexes posent parfois problème pour définir une visualisation aisée. Dans le cas des SPs, nous proposons d'utiliser des index-plot (Scherer, 2001) qui ont déjà été utilisés avec succès pour la visualisation de séquences. Le principe des index-plot est d'afficher une séquence représentant les activités comme des rectangles possédant une couleur par activité différente et avec une largeur proportionnelle à sa durée (*i.e.*  $stop(ac_i) - start(ac_i)$ ). Il est ainsi possible de visualiser facilement les SPs et d'effectuer une comparaison visuelle. La figure 1 présente une SP ( $sp_1 = \langle ac_1^{(1)}, ac_2^{(1)}, ac_3^{(1)} \rangle$ ) avec trois activités différentes.

## 3 Expériences

Vingt-quatre procédures (10 hommes, 14 femmes avec un âge médian de 52 ans) de chirurgie d'hernie lombaire ont été enregistrées dans le département de neurochirurgie de l'université de Leipzig en Allemagne. Une procédure peut être divisée en 3 principales étapes : l'approche du disque hernié, la discectomie comprenant la dissection et l'enlèvement du disque, et la phase de fermeture. Une étape d'hémostase peut parfois être nécessaire avant la phase de fermeture.

Cinq chirurgiens seniors et cinq chirurgiens juniors ont participé à cette étude ; Les chirurgiens seniors ayant déjà pratiqué plus de 100 retraits d'hernie, et les juniors ayant effectué au moins deux ans d'internat. Parmi les 24 procédures enregistrées, 12 ont été pratiquées par un chirurgien senior avec l'aide d'un chirurgien junior, et 12 ont été pratiquées par un chirurgien junior avec l'aide d'un chirurgien senior. Durant l'enregistrement des procédures effectuées par les chirurgiens juniors, seule la dernière étape a été pratiquée par le chirurgien junior sans l'aide du chirurgien senior. Ainsi, dans cette étude, nous nous sommes concentrés uniquement sur cette dernière étape pour pouvoir mieux discriminer les comportements entre chirurgien junior et senior. Les données ont été acquises à l'aide du logiciel Surgical Workflow Editor (Neumuth et al., 2006). Elles ont été enregistrées par un observateur humain, un neurochirurgien senior, qui a utilisé une tablette au cours de l'opération.

La Figure 2 présente le dendrogramme obtenu par l'application de l'algorithme de clustering hiérarchique ascendant (CHA) sur la phase de fermeture. L'index-plot représentant chacune

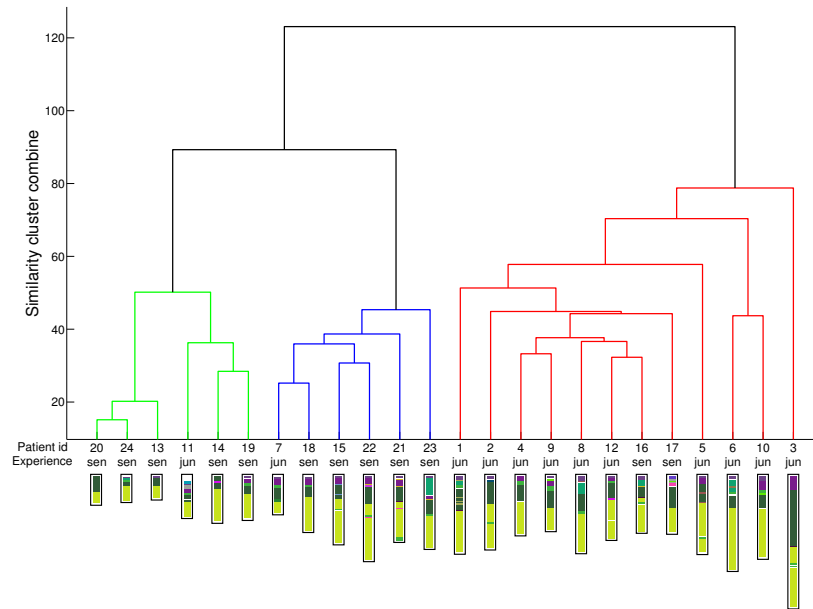


FIG. 2 – Dendrogramme résultant du clustering des 24 procédures considérées.

des SPs est affiché en dessous du dendrogramme. Trois clusters émergent visuellement de cette analyse. En coupant le dendrogramme pour obtenir 2 clusters, 12 chirurgies sont présentes dans chaque cluster. Le premier cluster (à gauche, en vert et bleu) contient 10 chirurgies pratiquées par des chirurgiens seniors et 2 par des chirurgiens juniors. Le second cluster (à droite, en rouge) quant à lui contient 10 chirurgies pratiquées par des chirurgiens juniors et 2 chirurgies pratiquées par des chirurgiens seniors. De plus, dans le premier cluster, deux sous-clusters sont présents (rouge et bleu), qui contiennent chacun 6 chirurgies. En conservant les deux clusters principaux, une précision de 83.33% peut être calculée en considérant que 20 chirurgies sur 24 sont bien classées.

## 4 Conclusion

Dans cet article, nous avons proposé une nouvelle métrique basée sur l'algorithme DTW pour la comparaison et l'évaluation de procédures chirurgicales. Cette approche permet de se concentrer sur l'analyse des différentes activités effectuées pendant une chirurgie ainsi que leur séquençage, et minimise l'importance des différences temporelles. Des résultats sur le clustering de procédures chirurgicales permettant d'identifier automatiquement le niveau d'expertise du chirurgien ont été présentés. Une des pistes que nous explorons actuellement s'intéresse à l'introduction de sémantique dans la mesure de similarité entre activités. En effet, la comparaison binaire utilisée pourrait être améliorée en prenant en compte la sémantique des actions, structures anatomiques et instruments, permettant ainsi une analyse plus précise.

## Clustering de séquences d'activités

Nous souhaitons également essayer de mettre en relation les résultats obtenus avec d'autres variables pré-opératoires (par exemple la difficulté de la chirurgie, le devenir du patient, etc.) afin d'expliquer les sous-clusters identifiés.

## Références

- Anderberg, M. (1973). Cluster analysis for applications. Technical report, Academic Press.
- Hiroaki, S. et S. Chiba (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 26, 43–49.
- Mackenzie, C., J. Ibbotson, C. Cao, et A. Lomax (2001). Hierarchical decomposition of laparoscopic surgery : a human factors approach to investigating the operating room environment. *Minimally Invasive Therapy and Allied Technologies* 10(3), 121–127.
- Neumuth, T., N. Durstewitz, M. Fischer, G. Strauß, A. Dietz, J. Meixensberger, P. Jannin, K. Cleary, H. Lemke, et O. Burgert (2006). Structured recording of intraoperative surgical workflows. In *SPIE Medical Imaging*, Volume 6145, pp. 61450A.
- Neumuth, T., P. Jannin, J. Schlomberg, J. Meixensberger, P. Wiedemann, et O. Burgert (2010). Analysis of surgical intervention populations using generic surgical process models. *International Journal of Computer Assisted Radiology and Surgery* 6, 59–71.
- Petitjean, F., A. Ketterlin, et P. Gançarski (2011). A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognition* 44, 678–693.
- Riffaud, L., T. Neumuth, X. Morandi, C. Trantakis, J. Meixensberger, O. Burgert, B. Trelhu, et P. Jannin (2010). Recording of surgical processes : a study comparing senior and junior neurosurgeons during lumbar disc herniation surgery. *Neurosurgery* 67, 325–332.
- Scherer, S. (2001). Early career patterns : A comparison of Great Britain and West Germany. *European Sociological Review* 17(2), 119.

## Summary

Towards the creation of new computer-assisted intervention systems, Surgical Process Models (SPMs) is an emerging concept used for analyzing and assessing surgical interventions. SPMs represent Surgical Process (SPs) which are formalized as symbolic structured descriptions of surgical interventions, using a pre-defined level of granularity and a dedicated terminology. In this context, an important challenge is the creation of new metrics for the comparison and the evaluation of SPs. Thus, the relations between these metrics and pre-operative data allow to classify surgeries and highlight specific information on the surgery itself and on the surgeon, such as his/her level of expertise. In this paper, we explore the automatic classification of a set of SPs based on the Dynamic Time Warping (DTW) algorithm. DTW allows to compute a distance between two SPs that focuses on the different types of activities performed during surgery and their sequencing, by minimizing time differences. Unsupervised classification experiments have shown that this approach was able to automatically identify groups of surgeons according to their level of expertise (senior and junior).