



HAL
open science

Modélisation par tranches horaires des habitudes de vie de personnes âgées pour l'aide à la prise de décision

Damien Brulin, Eric Campo, Clément Lejeune, Daniel Estève

► To cite this version:

Damien Brulin, Eric Campo, Clément Lejeune, Daniel Estève. Modélisation par tranches horaires des habitudes de vie de personnes âgées pour l'aide à la prise de décision. Journées d'Etude sur la TéléSanté, Sorbonne Universités, May 2019, Paris, France. hal-02161440

HAL Id: hal-02161440

<https://hal.science/hal-02161440v1>

Submitted on 20 Jun 2019

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.

Modélisation par tranches horaires des habitudes de vie de personnes âgées pour l'aide à la prise de décision

Damien Brulin¹, Eric Campo¹, Clément Lejeune¹, Daniel Estève¹

¹LAAS-CNRS, Université de Toulouse, CNRS, UT2J, Toulouse, France

dbrulin@laas.fr, ecampo@laas.fr

Abstract - Cet article explore de nouvelles étapes dans la mise en œuvre et l'exploitation de l'indicateur de présence afin d'identifier l'activité de déplacements de la personne et modéliser ses habitudes de vie. Cette modélisation s'appuie sur une mesure préalable des variations du taux de déplacement sur 24h dont l'analyse met en évidence un découpage possible des activités en tranches horaires, dont les bornes peuvent légèrement varier d'un jour à l'autre et d'une personne à l'autre. Sur cette base, nous montrons comment construire trois indicateurs nouveaux, en travaillant tranche horaire par tranche horaire, afin de détecter au plus tôt et de manière plus fiable des situations de danger ou dérives dans le comportement "normal" de la personne suivie. Ces indicateurs sont les horaires de début et de fin des tranches horaires, le taux de déplacements par tranche horaire, la durée de chaque tranche horaire. L'algorithme que nous proposons a été expérimenté sur des situations réelles pour montrer son usage et sa pertinence. Les résultats sont finalement intégrés dans une démarche plus ambitieuse de détection et d'aide à la décision automatique à travers la conception d'une interface web.

Keywords : modélisation du comportement, classification, maintien à domicile, aide à la décision

I. INTRODUCTION

L'intérêt de disposer de modèles du comportement humain s'est imposé dès le démarrage des travaux consacrés à la surveillance des personnes âgées à domicile : l'identification personnalisée des habitudes de vie est en effet apparue une base solide pour détecter tout changement pouvant témoigner d'un risque soudain pour la personne ou d'un changement plus lent de l'état de santé de celle-ci : événement inattendu, malaise, chute, diminution de la vitesse de marche, de l'activité physique... Ce concept a été largement appliqué par les auteurs de ce papier [1-2]. On le retrouve également dans toute la littérature spécialisée [3-7] comme base de détection des risques à domicile.

La mise en œuvre la plus immédiate est d'appliquer le concept directement avec des indicateurs de présence [4], en procédant par apprentissage. On observe, pendant la phase d'apprentissage, la présence du patient dans différents lieux stratégiques (salon, cuisine, chambre, entrée, ...) pour en dégager des probabilités de présence et écarts-types ; ces indicateurs, périodiquement réactualisés, vont permettre de définir un taux de présence habituel que l'on qualifiera de

« normal » pour la personne concernée et de détecter ainsi un défaut ou un excès de présence, une fois la phase d'apprentissage initiale terminée. On peut sur cette base imaginer un premier niveau de surveillance qui va émettre des alertes automatiques en cas de défauts ou d'excès de présences, grâce à des lieux de vie choisis pour leur capacité à témoigner d'habitudes fortement ancrées dans la vie quotidienne. Par exemple, le taux de présence dans la chambre permet d'obtenir l'heure du lever et du coucher qui sont généralement les mêmes d'un jour à l'autre.

Le suivi de personnes âgées à domicile utilisant pour seule instrumentation de l'habitat des capteurs de mouvement répartis dans différentes pièces est une approche maintenant bien connue et maîtrisée pour faire l'objet d'offres commerciales¹ que l'on retrouve classiquement dans des applications de télémonitoring, téléassistance pour l'aide au maintien à domicile. Cependant, il faut dans sa mise en œuvre trouver la procédure pour gérer au mieux les fausses alertes qui restent nombreuses.

Dans le cadre de cet article, nous nous intéressons aux méthodes et moyens permettant de réduire les fausses alertes en :

- s'appuyant davantage sur l'identification de la nature des activités du patient aux différentes heures du jour et de la nuit,
- exploitant plusieurs indicateurs et leurs convergences dans une logique de décision.
- proposant la possibilité de configurer des règles de décision propres à chaque personne et chaque tranche horaire à l'aide d'une interface web.

II. LE DECOUPAGE EN TRANCHE HORAIRE

Nous ne remettons pas en cause ici l'idée de s'appuyer sur le taux de présence pour construire une modélisation personnalisée des habitudes de vie à domicile : il reste un indicateur essentiel. La proposition est de valoriser cet indicateur en s'efforçant de le rendre plus significatif en le considérant dans son contexte d'activité.

¹ Telegrafik (<https://www.telegrafik.fr/>), Domosafety (<http://www.domo-safety.com/>), Bluelinea (<https://bluelinea.com/>)

A. Vers une amélioration de la modélisation des habitudes

Lors des premières réflexions, l'exploitation des trajectoires de déplacement dans l'habitat a été envisagée, sous la forme de trois indicateurs tirés des données de présence : quelles sont les trajectoires habituelles dans l'habitat ? Le patient est-il dans une trajectoire habituelle ? Quelle est sa vitesse de déplacement sur ce trajet ? Cette approche est intéressante dans la mesure où elle fait intervenir la vitesse de marche qui est un des cinq critères de la fragilité tel que présentée par Fried et al. [8], et qui est même considérée comme le sixième signe vital [9]. En effet, toute diminution lente de cette capacité à se déplacer, témoigne d'une certaine tendance au vieillissement du patient. Cette option a été explorée [10] sans aboutir à cause de la difficulté d'exploiter la notion de trajectoire « habituelle » qui suppose un point de départ et un point d'arrivée. On a pu toutefois observer que les caractéristiques de ces trajectoires sont très liées aux horaires.

C'est donc vers l'idée de relier la présence et les déplacements à la nature même des activités domestiques que nous nous sommes orientés pour améliorer la précision de nos modèles « habitudes de vie », faciliter la compréhension des observations et donc, finalement, réduire le taux de fausses alertes.

Pour concevoir cette approche, nous partons des données expérimentales collectées sur des personnes et traitées selon la logique d'analyse suivante : peut-on, par des études de classifications non supervisées, définir des tranches horaires où l'activité personnelle est reproductible de jour en jour, peut être caractérisée voire identifiée ? Si tel est le cas, il est alors envisageable de construire une modélisation des habitudes dans un contexte de plus grande cohérence des activités et de dégager des modèles plus précis et plus sensibles pour la détection des situations de danger. La problématique se résume alors à pouvoir définir ces tranches horaires, les identifier et en actualiser constamment les caractéristiques.

Pour démontrer la faisabilité de cette approche, nous nous appuyons sur des données réelles gracieusement fournies par l'entreprise Telegrafik.

B. Réflexion et conception d'un algorithme de découpage horaire automatisée

1) Présentation des données

Les données de présence concernent 6 personnes vivant seules à domicile pour lesquelles nous ne disposons que de peu

d'informations quant à leur âge, d'éventuelles pathologies communes. Pour chaque personne, cinq capteurs sont répartis dans l'habitat : 4 capteurs de mouvement respectivement dans la cuisine, le salon, l'entrée, et la chambre, et un capteur d'ouverture/fermeture de la porte d'entrée. Les capteurs de mouvement sont positionnés par des installateurs qualifiés afin de garantir une bonne visibilité (pas d'obstacle) et la meilleure couverture possible. Les capteurs choisis par l'entreprise Telegrafik sont « pet immune » permettant ainsi de minimiser les fausses détections dues à des animaux domestiques. Les zones sanitaires comme les toilettes ou la salle de bain ne sont en revanche pas couvertes. Les données brutes de présence sont formatées afin de ne conserver que les informations essentielles : date, numéro du jour dans la semaine, numéro du quart d'heure, pièce, présence/absence. Un exemple de données est illustré dans le Tableau 1.

Tableau 1 : Exemple de données de présence formatées

Echantillon	Date	Jour de la semaine	N° du quart d'heure	ID Client	Pièce	Etat
248686	2015-09-15 06:43:35	2	27	36	Kitchen	Presence
248739	2015-09-15 06:45:13	2	28	36	Entrance	Presence
248746	2015-09-15 06:45:19	2	28	36	Entrance	Presence
248763	2015-09-15 06:45:53	2	28	36	Entrance	Presence

2) Premières exploitations des données : taux de présence et nombre de détections

La première idée pour modéliser les habitudes est de chercher à découper une journée en plusieurs « tranches horaires » sur la base d'un indicateur d'activité observé sur une période significative. Lors d'une précédente étude [11], le nombre de détections par heure sur un échantillon de données de 30 jours a été choisi dans la mesure où il traduit le taux de présence dans une pièce mais également l'activité (mouvement dans une pièce et transition entre deux pièces). Le résultat de cette classification non supervisée (le nombre de tranches horaires n'étant pas fixé a priori) est illustré par la Figure 1.

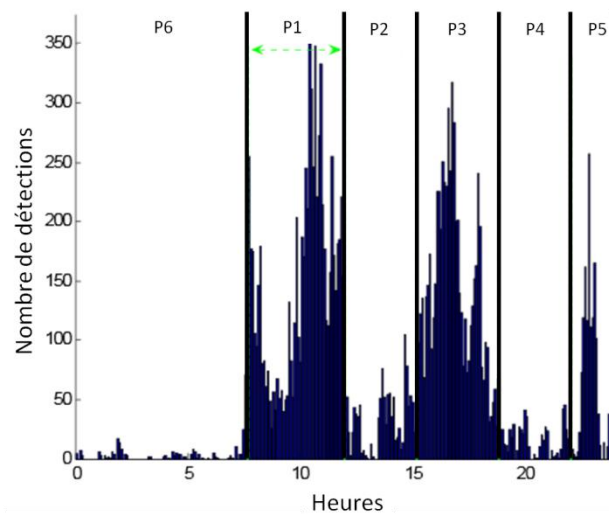


Figure 1. Classification temporelle sur 24h sur une période de 30 jours

Cette classification temporelle met en évidence 6 périodes caractéristiques du comportement à domicile au quotidien de cette personne. On constate une activité faible la nuit, entre minuit et 7h30 du matin, forte la matinée, moyenne pendant la période du repas, forte l'après midi et de nouveau plus faible dans la soirée avec des pics au moment du coucher. Ici la géographie des lieux et la disposition des capteurs sont connues. Ces caractéristiques sont bien sûr indépendantes des capteurs installés qui donnent une simple indication binaire présent/pas présent. Pour cette étude, 12 capteurs ont été

utilisés pour une superficie de 35 m², cela permet une localisation assez précise de la personne.

Une étape supplémentaire de caractérisation est de relier l'activité à chaque site de détection, c'est à dire, le plus souvent à une pièce de l'habitat : notre indicateur de présence est quantifié en taux de présence par rapport à la présence totale pour normaliser les données et les manipuler plus facilement. Nous avons alors utilisé les données Telegrafik pour réaliser cette nouvelle classification qui est illustrée à la Figure 2. Elle montre le taux de présence par site (ici chambre, entrée, cuisine, salon) sur 24h et sur une semaine d'observation.

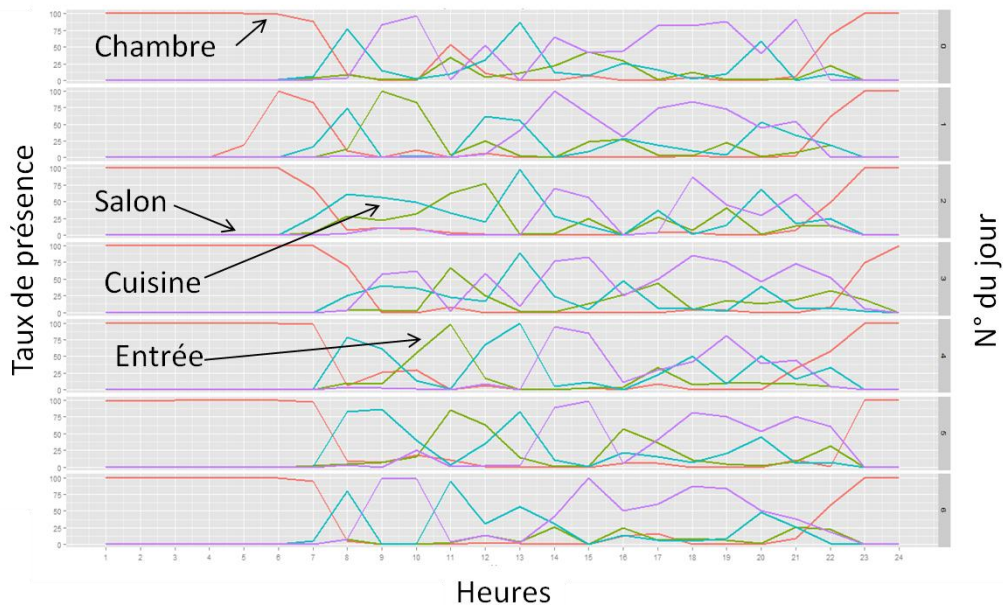


Figure 2. Taux de présence par heure sur une semaine

Pour cette personne, on constate tout d'abord que certaines pièces (chambre, cuisine) possèdent un profil de présence similaire sur tous les jours de la semaine. De plus, on peut distinguer que certaines transitions entre deux pièces se font à des horaires très proches (passage chambre/cuisine vers 7h30 par exemple). Enfin, il est possible de retrouver les activités de préparation des repas en observant les courbes cuisine et salon. Ainsi, le taux de présence semble être un indicateur pertinent pour représenter les habitudes comportementales d'une personne (ces constatations ayant été observées chez d'autres personnes). Si l'on s'intéresse maintenant au nombre de détections (activités) en sommant toutes les détections dans toutes les pièces sur une période de trois semaines, on obtient pour la même personne les résultats de la Figure 3.

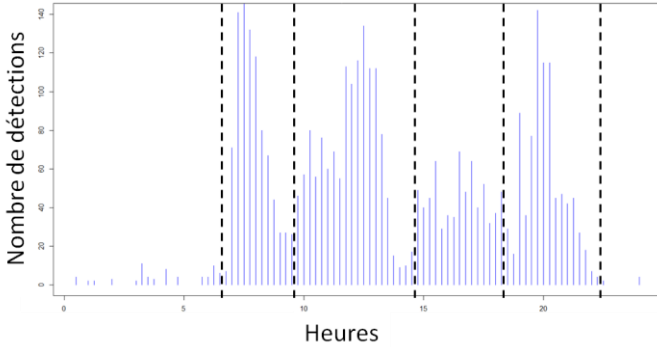


Figure 3. Nombre de détections par 1/4h sur 3 semaines

Des tranches horaires sont alors davantage visibles. Ce graphe montre, qu'avec ce système de capteurs, la journée peut se décomposer en 6 tranches horaires. Cette mise en évidence nécessite cependant 3 semaines de données, ce qui est assez conséquent.

3) Algorithme de détermination des tranches horaires

L'objectif est ici de traiter les données de présence et d'activité pour obtenir un modèle empirique des tranches horaires. La première idée est d'estimer le temps d'apprentissage nécessaire pour voir se stabiliser ces tranches par traitement direct des données. Pour cette évaluation, nous avons cumulé des données sur trois semaines comme l'illustre la Figure 3.

Si l'on veut par exemple, chaque jour, réajuster les tranches horaires (la durée et les instants de début/fin), il est difficilement concevable d'utiliser une fenêtre glissante sur 21 jours. Cela implique de choisir des tranches horaires « quasi fixes » (l'ajustement des tranches horaires, s'il y a changement progressif des habitudes, serait trop lent).

La démarche que nous avons choisi d'appliquer est de construire un modèle durant une période d'apprentissage de 21 jours, caractérisé par l'activité et la présence, selon 6 tranches horaires, et d'en extraire trois indicateurs :

- la durée des tranches horaires,
- les horaires de début et de fin,
- le niveau d'activité moyen dans chaque tranche horaire.

Une fois la phase d'apprentissage terminée, on corrige quotidiennement chaque indicateur en pondérant la valeur moyenne apprise et la valeur nouvelle courante ; on dispose ainsi d'une actualisation journalière et de la possibilité de s'appuyer sur la valeur régulièrement actualisée pour ouvrir la possibilité d'une surveillance « longitudinale » et ainsi de comparer des graphes de même type sur plusieurs mois. Ceci constitue en effet une demande systématique du monde médical qui souhaite pouvoir évaluer de manière objective l'effet de ses prescriptions.

Suite à ces considérations, nous avons appliqué la démarche suivante :

- Tout d'abord, nous définissons une façon de découper les journées en nombre de tranches horaires fixes, en choisissant un indicateur et un seul pour diminuer la complexité,
- Dans un second temps, nous définissons une procédure journalière d'apprentissage de ces tranches horaires,
- enfin, nous surveillons les dispersions : les valeurs des indicateurs de suivi choisis seront encadrées par une valeur maximum et une valeur minimum, résultant également d'une procédure d'apprentissage.

Il est nécessaire d'évaluer la sensibilité du modèle afin d'aboutir à un niveau de fausses alertes acceptable pour la personne, en resserrant ou relaxant les valeurs d'encadrement des indicateurs (maximum et minimum appris).

III. MISE EN ŒUVRE DU DECOUPAGE EN TRANCHES HORAIRES

Dans cette partie, on présente le raisonnement étape par étape illustré par des équations sachant que l'intégralité du codage a été réalisée en langage R.

A. Indicateur du taux de présence

Soit Δ_{pres} , le temps de présence dans la pièce r , au quart d'heure $4t-q$ et $\tau_{pres}(t,r)$, le taux de présence à l'heure t , dans la pièce r .

$$0 \leq \Delta_{pres}(4t - q, r) \leq 15 \min(1)$$

$$\forall t \in \llbracket 1, 24 \rrbracket, \tau_{pres}(t, r) = \frac{100}{60} \sum_{q=0}^3 \Delta_{pres}(4t - q, r) \quad (2)$$

Rappelons qu'il est nécessaire de prendre en compte la possibilité que la personne soit présente mais non détectée dans la mesure où toutes les pièces de l'habitat ne sont pas couvertes (toilettes, salle de bain...). De plus, la position des capteurs peut également être source d'erreurs (la personne est peut-être présente mais pas vue par le capteur). L'algorithme de calcul de présence prend en entrée les données formatées pour le jour considéré et délivre en sortie le temps de présence par pièce et par quart d'heure. Pour chaque quart d'heure, deux situations sont envisagées :

- personne présente mais inactive : création de présence au début du/des quarts d'heure non détectés dans la pièce de la dernière détection.
- personne présente et quart d'heure présent dans les données : création du quart d'heure suivant si inexistant.

B. Calcul des tranches horaires

L'objectif est de déterminer les « instants réels » de fin de tranche horaire. Ils se définissent comme des « points de transition » et se base sur des critères empiriques simples. Comme indicateur privilégié nous avons choisi le taux de présence pour réaliser la division en 6 tranches horaires et aboutir à un modèle empirique personnalisé. L'exécution en temps réel intervient à la fin de chaque jour (estimation) et les données sont intégrées dans la procédure d'apprentissage. On fixe au départ, de manière empirique, les tranches horaires T_k (instants de début/fin), puis on identifie la pièce R_k la plus fréquentée, pour chaque tranche horaire et on estime les « instants réels » de fin de tranche (voir Tableau 2).

Tableau 2 : Estimation de la fin et pièce attribuée à chaque tranche

	deb	fin	room	fin.estim
T6	1	7	BEDROOM	7.067168
T1	7	12	KITCHEN	10.484733
T2	12	14	KITCHEN	13.495745
T3	14	17	LIVINGROOM	16.494167
T4	17	19	LIVINGROOM	18.998171
T5	19	21	LIVINGROOM	21.918875

Pour obtenir ces estimations, les données sont lissées par une procédure d'interpolation afin de rendre plus fonctionnels les taux de présence en fonction des heures. Il existe de nombreuses méthodes d'interpolation (spline cubique, circulaire, B-spline, polynômes...), dans notre cas, on a choisi les splines cubiques pour la simplicité de mise en œuvre.

Pour repérer les passages d'une tranche horaire à la suivante, nous avons été amenés à appliquer deux critères différents :

(1) Lorsque la pièce ayant le taux d'occupation principal, diffère d'une tranche horaire à une autre consécutive, le critère choisi est de détecter « l'inversion de présence » c'est-à-dire de déterminer l'instant \hat{t}_k le plus proche de t_k (la fin théorique de T_k , fixée à l'étape initiale) où le taux de présence dans R_{k+1} devient supérieure à celui dans R_k entre le début de T_k et la fin de T_{k+1} . Graphiquement, il s'agit simplement du point d'intersection entre les deux courbes (voir Figure 4).

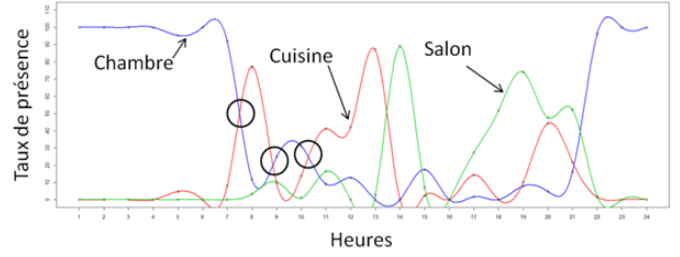


Figure 4. Exemple de transitions entre T6 (chambre) et T1 (cuisine)

(2) Dans ce cas, la pièce occupée reste la même pour les deux tranches horaires consécutives. On souhaite partager la courbe de la pièce concernée en deux parties de manière à produire une frontière temporelle \hat{t}_k proche du t_k fixé initialement.

La méthode employée se base sur un algorithme de classification non supervisée, il s'agit de l'algorithme des k-médoïdes. Cet algorithme est très semblable à celui des k-moyennes dont l'objectif est de déterminer k points moyens m_i pour lesquelles les dispersions intra-classes sont minimales au sens de la distance euclidienne. k est choisi arbitrairement, dans notre cas on impose $k = 2$.

Étant donné un ensemble d'observations $\{x_1 \dots x_n\}$ de R^q , la classe C_i est représentée par son centre m_i (moyenne de la classe sur les q modalités des observations) :

$$C = \{C_1 \dots C_{k < n}\} = \arg_c \min \sum_{i=1}^k \sum_{x_j \in C_i} \|x_j - m_i\|^2 \quad (3)$$

$$m_i = \frac{1}{n_i} \sum_{x_j \in C_i} x_j \quad (4)$$

La différence avec les k-médoïdes est qu'une classe est représentée par un médoïde M_i , c'est-à-dire un point des données dont la moyenne des dissimilarités à tous les points de la classe est minimale au sens de la distance euclidienne. La dissimilarité entre deux observations est ici la distance euclidienne.

$$diss(x_1, x_2) = \|x_1 - x_2\|^2 \quad (5)$$

La dissimilarité moyenne d'une classe C_i est la distance moyenne entre tous les individus de cette classe à son médoïde M_i .

$$diss(C_i) = \frac{1}{n_i} \sum_{x_j \in C_i} \|x_j - M_i\|^2 \quad (6)$$

$$M_i = \arg_{x_i \in \{x_1 \dots x_n\}} \min diss(C_i) \quad (7)$$

La détermination des classes se fait de la même façon qu'avec les k-moyennes en remplaçant m_i par M_i dans Eq. 3.

A l'étape initiale, l'algorithme choisit k points aléatoirement qui représentent les classes. Il affecte les individus à chaque classe la plus proche. Puis il calcule les centres des classes (moyennes ou médoïdes des Eq. 4 et Eq. 7). Ces deux dernières

étapes sont répétées jusqu'à convergence, c'est-à-dire tant que la valeur du critère (Eq. 3) est inférieure à celle de l'étape précédente.

L'algorithme des k-médoïdes est moins sensible aux données aberrantes que l'algorithme des k-moyennes mais son temps d'exécution est plus grand en raison du calcul médoïde plus long que celui de la moyenne. Pour notre cas, cela n'est pas un problème car la taille des données de présence sur deux tranches horaires (points des courbes interpolées) est assez petite (~ 50 points).

Une fois les 2 médoïdes obtenus, on calcule la moyenne afin d'obtenir le barycentre des deux classes qui correspond au \hat{t}_k recherché.

$$\hat{t}_k = \frac{(M_1 + M_2)}{2} \quad (8)$$

Une tranche horaire estimée le jour d , $\hat{T}_{k,d}$, comporte alors un début (qui est en fait l'estimation de la fin de la tranche horaire précédente) et une fin estimée.

$$\hat{T}_{k,d} = [\hat{t}_{k-1}; \hat{t}_k] \quad (9)$$

C. Déviance des tranches horaires

Les estimations des tranches horaires restent fortement liées aux taux de présence. Les variations du début et/ou de la fin d'une tranche peuvent signifier une modification du comportement habituel et constituent donc un second indicateur pertinent.

Tout d'abord, pour chaque tranche k , on définit deux transitions moyennes (début = 1 et fin = 2) sur $n = 3$ jours, $\bar{T}_{k,1,d-1}$ et $\bar{T}_{k,2,d-1}$ au jour $d - 1$ calculées par moyenne glissante sur les estimations $\hat{T}_{k,1,d-1} \dots \hat{T}_{k,1,d-n}$ (voir équation x).

$$\bar{T}_{k,1,d-1} = \frac{1}{n} \sum_{q=d-n}^{d-1} \hat{T}_{k,1,q} \quad (10)$$

On procède de même pour les écarts-types associés à $\bar{T}_{k,1,d-1}$ et $\bar{T}_{k,2,d-1}$.

$$\bar{\sigma}_{k,1,d-1} = \sqrt{\frac{1}{n-1} \sum_{q=d-n}^{d-1} (\hat{T}_{k,1,q} - \bar{T}_{k,1,d-1})^2} \quad (11)$$

Etant donné les estimations $\hat{T}_{k,1,d}$ et $\hat{T}_{k,2,d}$ d'une tranche, l'objectif est de construire un intervalle de tolérance pour le début et un intervalle pour la fin afin de quantifier les écarts $D_{k,1,d}$ et $D_{k,2,d}$ de ces estimations par rapport à leur moyenne et écart-type respectif à $d - 1$.

En d'autres termes, on considère que la personne ne sort pas de ses habitudes si l'estimation $\hat{T}_{k,1,d}$ est dans l'intervalle $[\bar{T}_{k,1,d-1} - \bar{\sigma}_{k,1,d-1}; \bar{T}_{k,1,d-1} + \bar{\sigma}_{k,1,d-1}]$. Les écarts respectent donc l'équation suivante :

$$D_{k,i,d} = \frac{\hat{T}_{k,i,d} - \bar{T}_{k,i,d-1}}{2\bar{\sigma}_{k,i,d-1}} \quad (12)$$

L'intervalle de tolérance pour ces écarts est donc $[-0.5; 0.5]$. On regroupe ces écarts en définissant l'écart absolu total moyen.

$$D_{k,d} = \frac{|D_{k,1,d}| + |D_{k,2,d}|}{2} \times 100 \quad (13)$$

D. Mise en œuvre de l'apprentissage

L'apprentissage que nous avons conçu vise à suivre le comportement au plus près, c'est-à-dire sans « trop » atténuer ou exagérer la variation du comportement dans le temps. Il servira alors de référence pour détecter des changements d'habitude.

L'idée est à chaque fois d'apprendre par le biais d'opérateurs statistiques simples sur une fenêtre glissante de taille fixe. Les valeurs apprises sont annotées d'un « * ».

1) Tranches horaires

L'apprentissage de la fin (indice 2) d'une tranche horaire $T^*_{k,2,d+1}$ pour le jour suivant $d + 1$ se base sur l'estimation et l'apprentissage du jour courant et du passé selon la loi suivante.

$$T^*_{k,2,d+1} = T^*_{k,2,d} + \lambda \left(\hat{T}_{k,2,d} - \frac{1}{n} \sum_{q=d-n+1}^d T^*_{k,2,q} \right) \quad (14)$$

Avec n la taille de la fenêtre glissante et $\lambda \in]0; 1[$. En réalisant un apprentissage sur des données dont on connaît la vérité terrain, il apparaît qu'un λ compris entre 0,4 et 0,6 fournit la meilleure estimation.

2) Taux de présence

L'apprentissage de la présence est fait par tranche horaire pour lesquelles on calculera un minimum moyen et un maximum moyen glissants (pour la pièce principale de la tranche). Pour une tranche T^*_k apprise, les deux lois appliquées sont les suivantes :

$$\max^*_{T^*_{k,d+1}}(pres) = \frac{1}{n} \sum_{q=d-n+1}^d \max_{t \in T^*_{k,q}} \tau_{pres}(t, R_k)_q \quad (15)$$

$$\min^*_{T^*_{k,d+1}}(pres) = \frac{1}{n} \sum_{q=d-n+1}^d \min_{t \in T^*_{k,q}} \tau_{pres}(t, R_k)_q \quad (16)$$

La Figure 5 montre les résultats obtenus pour une fenêtre glissante de 4 jours sur une période de 21 jours. En rouge on retrouve l'apprentissage du maximum moyen, en vert celui du minimum, en noir l'estimation et en bleu la moyenne glissante, cette dernière étant ajoutée à titre indicatif.

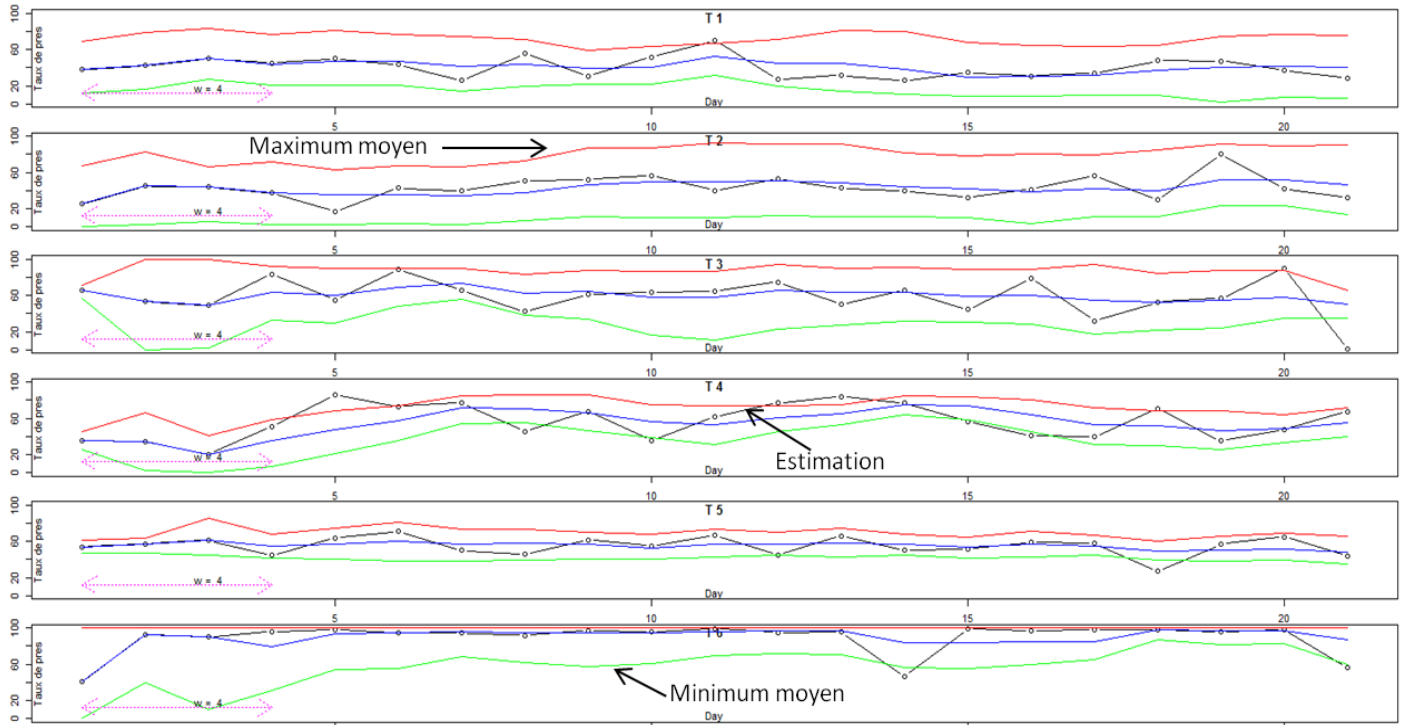


Figure 5. Apprentissage de la présence sur 21 jours avec une fenêtre glissante de 4 jours pour chaque tranche horaire

3) Détection des situations inhabituelles

La détection de taux de présence inhabituel se base sur l'écart entre le taux de présence courant à l'instant t dans la pièce principale R_k et les valeurs apprises de l'indicateur pour $t \in T_{k,d}^*$.

$$E(\tau_{\text{pres}}(t, R_k)) = \frac{\max_{T_{k,d}^*}^*(\text{pres}) - \tau_{\text{pres}}(t, R_k)}{\max_{T_{k,d}^*}^*(\text{pres}) - \min_{T_{k,d}^*}^*(\text{pres})} \times 100 \quad (17)$$

4) Indicateur de déviance de la tranche horaire

De la même manière que pour le taux de présence, l'apprentissage de l'indicateur de déviance des tranches horaires se fait par encadrement d'un minimum et d'un maximum moyens glissants sur une fenêtre de taille fixe n . La formule diffère puisqu'ici une seule valeur de déviance existe par tranche tandis que pour la présence, il s'agissait de prendre les valeurs extrêmes par tranche horaire, c'est-à-dire de moyenner les n derniers taux de présence maxi (ou minimum) dans la pièce principale et dans la tranche horaire considérée. Ici, on moyenne la déviance maximale (ou minimale) des n derniers jours avec celle apprise de la veille. De par leur nature, la différence entre la présence et la déviance apprises réside en fait dans l'échelle de temps sur lequel on calcule ces valeurs extrêmes.

$$\max_{T_{k,d+1}^*}^*(dev) = \frac{\max_{q \in \llbracket d-n+1; d \rrbracket} E_{k,d} + \max_{T_{k,d}^*}^*(dev)}{2} \quad (18)$$

$$\min_{T_{k,d+1}^*}^*(dev) = \frac{\min_{q \in \llbracket d-n+1; d \rrbracket} E_{k,d} + \min_{T_{k,d}^*}^*(dev)}{2} \quad (19)$$

Pour les n premiers jours on initialise à 0. Contrairement à la présence, la déviance n'est pas bornée, on décide de rejeter celles supérieures à 1 pour l'apprentissage. La déviance apprise est donc comprise entre 0 et 1. Autrement, cela « gonflerait » inutilement \max^* .

$$E(D_{k,d}) = \frac{\max_{T_{k,d}^*}^*(dev) - D_{k,d}}{\max_{T_{k,d}^*}^*(dev) - \min_{T_{k,d}^*}^*(dev)} \times 100 \quad (20)$$

IV. L'AIDE A LA PRISE DE DECISION

L'écart sert à évaluer le niveau de danger de la situation. Ce niveau sera qualifié en fonction de la valeur de cet écart par rapport à un seuil s_1 et un autre seuil s_2 . Si on considère trois niveaux d'alerte « vert », « orange » et « rouge », la décision pourra par exemple suivre le schéma suivant :

- rouge si écart < 0 ou $> s_2$
- orange si écart entre s_1 et s_2
- vert si écart $< s_1$

Les valeurs des seuils d'alerte peuvent être dans un premier temps fixées (choisies par l'utilisateur de l'application temps-réel) et pourront être ajustées dans le temps s'il y a trop d'alertes par exemple (cet aspect est un point de perspectives de nos travaux et devrait se faire dans l'idéal automatiquement sans intervention humaine).

La mesure des écarts à la normalité sur un seul indicateur présente des limites évidentes qu'il faut maîtriser par deux mécanismes essentiels :

- La multiplication d'indicateurs de manière à circonscrire plus précisément l'origine des écarts à la normale constatés,

- La prise en compte du contexte et de l'environnement de la personne surveillée : il est clair que pour une personne sujette aux malaises, il convient d'en tenir compte dans les critères et les règles de détection du danger et de décision sur les mesures à prendre.

Nous préconisons de travailler à deux niveaux d'agrégation des indicateurs :

- Un niveau technique qui vise, par la multiplication des indicateurs, de s'assurer de la mesure de l'écart à la normalité comportementale,
- Un niveau d'aide à la décision qui vise à aider les opérateurs à évaluer le danger qui est associé à l'alarme en fonction de la personne, de son environnement et des moyens d'intervention qu'il a à sa disposition.

Ainsi, ces indicateurs font partie des entrées d'une première table de détection et peuvent être combinés à travers des règles simples (de type SI...ET SI... ALORS...) afin de délivrer un premier niveau d'alerte. Par exemple, un écart de présence ORANGE associé à un écart de déviance ROUGE donnera un niveau de danger ROUGE. Cette détection est nécessaire mais pas suffisante pour déclencher une intervention notamment sur des situations de type ORANGE. Ainsi, une seconde table de décision, permettra, selon des règles définies en fonction de la sortie de la première table de détection, du contexte de vie, du profil de la personne, de son état de santé..., de proposer l'intervention la plus pertinente et adaptée à cette personne.

Une application nommée Vigilesoft a été pensée et développée afin de permettre une configuration aisée, par l'utilisateur final lui-même (surveillant, médecin, ...) de règles expertes pour les deux tables : détection et décision. Cette application a pour objectif, une fois la configuration des tables effectuée, de permettre des simulations sur des jeux de données existantes mais surtout une surveillance temps réel des personnes et un suivi longitudinal indispensable aux médecins pour la détection de dérive lente du comportement. Cette plateforme logicielle s'inscrit pleinement dans la poursuite de nos travaux sur la détection de fragilité de systèmes complexes (homme, structure, environnement) et a donc été pensée pour être indépendante du système suivi.

V. CONCLUSION ET PERSPECTIVES

Le capteur de présence est un capteur essentiel dans la surveillance des personnes âgées à domicile. Il permet de calculer un taux de présence variable dont on peut comparer, à tout instant, la valeur moyenne et la valeur instantanée comme base de détection d'une situation anormale. La mise en œuvre de ce concept est aujourd'hui généralisée mais donne des résultats insuffisants de par le nombre élevé de fausses alertes.

Dans cet article, nous explorons l'objectif d'améliorer la précision de la modélisation des habitudes de vie en construisant une modélisation par « tranches horaires », chaque tranche horaire étant caractéristique d'une activité domestique dont on peut penser qu'elle est nécessairement reproduite de jour en jour. Nous pensons ainsi pouvoir construire des modèles plus précis et plus sensibles permettant de détecter des situations qui s'écartent du comportement habituel.

Partant de données empiriques issues de plusieurs personnes, nous montrons que l'on peut imaginer un découpage automatique en tranches horaires où les activités sont relativement caractéristiques et reproductibles. De manière à exploiter pratiquement ce découpage, nous le supposons fixe (6 tranches horaires) et calculons, jour après jour, les dates de début et de fin de la tranche horaire. Nous définissons ainsi plusieurs indicateurs : date de début et de fin de la tranche horaire, taux d'activité dans la tranche horaire, taux de présence par pièce dans la tranche horaire. Nous introduisons, dans cet article, les fondements que nous utilisons pour associer dans un premier temps, dans une table de détection, plusieurs indicateurs par des jeux de règles ouvertes permettant des ajustements manuels et une compréhension plus aisée de la situation de terrain. Cette détection d'un écart par rapport à la normalité (telle qu'elle a été modélisée) est replacée dans le contexte de vie de la personne (antécédents, environnement...) afin de parvenir à une décision d'intervention la plus pertinente possible. Cela a fait l'objet de la création d'une application qui est encore en développement actuellement et qui constituera un outil essentiel dans la poursuite et la valorisation de nos travaux.

Notre étude confirme que la mesure de présence par des capteurs infrarouges de faible coût, reste une voie essentielle dans la conception des systèmes de surveillance à domicile. Cette étude reste toutefois perfectible. D'une part, une phase préliminaire de nettoyage des données brutes de mouvement semble indispensable afin d'éliminer les données incohérentes dues à des perturbations de l'environnement (animaux de compagnie, flux d'air chaud,...). D'autre part, l'avenir est de pouvoir, par la multiplication des capteurs et leur positionnement optimal dans la prise d'information, accéder à des découpages plus détaillés et plus précis donnant accès à la nature même des différentes activités domestiques ; une modélisation encore plus précise de la normalité est possible qui permettra de limiter largement les fausses alarmes qui font obstacle au développement en masse de ces technologies.

REMERCIEMENTS

Les auteurs de l'article remercient la Région Midi-Pyrénées pour le financement (programme READYNOV) du projet TELEPASS, dans lequel s'inscrivent ces travaux. Nos remerciements vont également à l'entreprise Telegrafik pour la mise à disposition des données et les échanges enrichissants qui ont permis d'engager de nouvelles réflexions et développements.

RÉFÉRENCES

- [1] Y. Charlon, W. Bourennane, F. Bettahar, E. Campo, Activity monitoring system for elderly in a context of smart home, IRBM, Volume 34, Issue 1, 2013, Pages 60-63.
- [2] W. Bourennane, Y. Charlon, F. Bettahar, E. Campo, D. Esteve, Homecare monitoring system: A technical proposal for the safety of the elderly experimented in an Alzheimer's care unit, IRBM, Volume 34, Issue 2, 2013, Pages 92-100.
- [3] P. Barsocchi, M. G.C.A. Cimino, E. Ferro, A. Lazzeri, F. Palumbo, G. Vaglini, Monitoring elderly behavior via indoor position-based stigmergy, Pervasive and Mobile Computing, Volume 23, 2015, Pages 26-42.

- [4] N. Zouba, F. Bremond, M. Thonnat, An Activity Monitoring System for Real Elderly at Home: Validation Study, *2010 7th IEEE International Conference on Advanced Video and Signal Based Surveillance*, Boston, MA, 2010, pp. 278-285.
- [5] A.W. Mitas, M. Rudzki, M. Skotnicka, P. Lubina, Activity Monitoring of the Elderly for Telecare Systems - Review. In: Piętko E., Kawa J., Wieclawek W. (eds) *Information Technologies in Biomedicine*, Volume 4. *Advances in Intelligent Systems and Computing*, vol 284. Springer, Cham, 2014.
- [6] N. Noury *et al.*, Monitoring behavior in home using a smart fall sensor and position sensors, *1st Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology. Proceedings (Cat. No.00EX451)*, Lyon, France, 2000, pp. 607-610.
- [7] H. Medjahed, D. Istrate, J. Boudy, J. Baldinger, B. Dorizzi, A pervasive multi-sensor data fusion for smart home healthcare monitoring, *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, Taipei, 2011, pp. 1466-1473.
- [8] Fried LP, Tangen CM, Walston J, Newman AB, Hirsch, C, Gottdiener, J, Seeman, T, Tracy, R, Kop, WJ, Burke, G, McBurnie, MA (Mar 2001). "Frailty in older adults: evidence for a phenotype". *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*. 56 (3): M146–56.
- [9] S. Fritz, M. Lusardi. White Paper: Walking Speed: the Sixth Vital Sign. *Journal of Geriatric Physical Therapy*, 32(2), p. 2-5, 2009.
- [10] E. Campo, M. Chan, W. Bourennane, D. Estève, Behaviour monitoring of the elderly by trajectories analysis, *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Buenos Aires, 2010, pp. 2230-2233.
- [11] S. Bonhomme, E. Campo, D. Estève, J. Guennec, PROSAFE-extended, a telemedicine platform to contribute to medical diagnosis. *Journal of Telemedicine and Telecare*, 14(3), 116–119, 2018.