UQAC
Université du Québec
à Chicoutimi

DÉVELOPPEMENT D'UNE APPROCHE ALGORITHMIQUE POUR LA DÉTECTION AUTOMATIQUE DES CHANGEMENTS DE NIVEAU D'ACTIVITÉ D'UNE PERSONNE ATTEINTE DE DYSTROPHIE MYOTONIQUE DE TYPE 1

PAR PATRICK LAPOINTE

MÉMOIRE PRÉSENTÉ À L'UNIVERSITÉ DU QUÉBEC À CHICOUTIMI
COMME EXIGENCE PARTIELLE EN VUE DE L'OBTENTION DU GRADE DE
MAÎTRE ÈS SCIENCES EN INFORMATIQUE

QUÉBEC, CANADA

© PATRICK LAPOINTE, 2021

RÉSUMÉ

La Dystrophie myotonique de type 1 (DM1) est une maladie neuromusculaire incurable qui, entre autres, cause une perte de capacité musculaire au fil des années. Il a été démontré qu'un programme d'entraînement prescrit par un spécialiste peut être bénéfique pour les personnes atteintes de cette maladie. D'ailleurs, le projet de maîtrise présenté dans ce mémoire se greffe au projet Acti-DM1 ayant pour but le développement d'une solution technologique d'assistance à domicile pour aider ce type de clientèle à réaliser un programme d'entraînement à domicile.

La technologie consiste principalement en un bracelet à centrale inertielle enregistrant différentes activités physiques (se lever, s'asseoir, marcher, courir) en temps réel connectée à une unité centrale de traitement possédant un haut-parleur. L'unité centrale de traitement s'occupe de recueillir les données du bracelet et de gérer le programme d'entraînement à domicile prescrit en effectuant des rappels quotidiens et en guidant le participant lors de ses exercices.

Au cours du projet Acti-DM1, la technologie d'assistance développée a été déployée dans 10 résidences de personnes atteintes de DM1. Lors de cette cueillette de données en situation réelle d'une durée de 10 semaines, beaucoup de données ont été recueillies. Le projet de recherche dont il est question dans ce mémoire a pour objectif de traiter ces données et de détecter les changements comportementaux au niveau de l'activité physique enregistrée par le bracelet à centrale inertielle au cours de la cueillette. Pour y arriver, des prétraitements importants ont dû être effectués sur les données brutes afin d'extraire l'information utile. Ensuite, plusieurs algorithmes constituant l'état de l'art au niveau de la détection automatique de changements comportementaux ont été étudiés, implémentés et appliqués.

TABLE DES MATIÈRES

RÉSUMÉ ii
LISTE DES TABLEAUX vi
LISTE DES FIGURES vii
LISTE DES ABRÉVIATIONS xii
DÉDICACE xiii
REMERCIEMENTS xiv
CHAPITRE I – INTRODUCTION 1
1.1 CONTEXTE 1
1.2 PROBLÉMATIQUE 3
1.2.1 TERMINOLOGIE ET CONCEPTS 3
1.2.2 HYPOTHÈSES 9
1.2.3 QUESTION DE RECHERCHE 11
1.3 CONTRIBUTION 11
1.4 MÉTHODOLOGIE 12
1.5 ORGANISATION DU DOCUMENT 13
CHAPITRE II – ÉTAT DE L'ART 15
2.1 ALGORITHMES 15
2.1.1 MÉTHODES SUPERVISÉES 16
2.1.2 MÉTHODES NON SUPERVISÉES 18
2.1.3 RULSIF 20
2.1.4 PCAR 23
2.1.5 SW-PCAR 31
2.2 CONTEXTES D'APPLICATION 34
2.2.1 HABITATS INTELLIGENTS 34

2.2.2 DÉTECTION DE CHANGEMENT AU NIVEAU DE L'ACTIVITÉ PHYSIQUE . . . 37
2.3 DISCUSSION . . . 40
CHAPITRE III – CUEILLETTE DE DONNÉES . . . 42
3.1 DISPOSITIFS . . . 42
3.1.1 BRACELET À CENTRALE INERTIELLE . . . 43
3.1.2 UNITÉ CENTRALE DE TRAITEMENT . . . 50
3.1.3 CAPTEUR DE VITESSE . . . 51
3.1.4 PIR+ . . . 53
3.2 DÉROULEMENT DE LA CUEILLETTE . . . 56
3.2.1 PRÉPARATION . . . 57
3.2.2 EXÉCUTION . . . 58
3.2.3 FINALISATION . . . 59
CHAPITRE IV – TRAITEMENT DES DONNÉES . . . 61
4.1 CONVERSION DES DONNÉES BRUTES . . . 61
4.1.1 REPRÉSENTATION INITIALE . . . 62
4.1.2 REPRÉSENTATION FINALE . . . 64
4.1.3 PROCESSUS DE CONVERSION . . . 67
4.2 EXTRACTION DE L'INFORMATION . . . 70
4.2.1 ÉTAPE 1 : CHARGEMENT DES DONNÉES FORMATÉES . . . 71
4.2.2 ÉTAPE 2 : PRÉTRAITEMENT - GESTION DES DÉCONNEXIONS . . . 72
4.2.3 ÉTAPE 3 : CHARGEMENT DES DONNÉES D'ENTRAÎNEMENT . . . 76
4.2.4 ÉTAPE 4 : COMPILATION DE L'INFORMATION PAR PÉRIODE . . . 78
4.3 SOMMAIRE . . . 89
CHAPITRE V – DÉTECTION DE CHANGEMENT COMPORTEMENTAL . . . 90
5.1 MÉTHODE . . . 90
5.1.1 MODES DE COMPARAISON . . . 91
5.1.2 CHARGEMENT DES CARACTÉRISTIQUES . . . 92

5.1.3 ALGORITHMES 94
5.1.4 SORTIE 99
5.2 RÉSULTATS DE L'ANALYSE DES CHANGEMENTS COMPORTEMENTAUX 102
5.2.1 PARTICIPANT 481 103
5.2.2 PARTICIPANT 907 107
5.2.3 PARTICIPANT 1453 112
5.2.4 PARTICIPANT 1692 113
5.2.5 PARTICIPANT 2002 118
5.2.6 PARTICIPANT 2182 122
5.3 RÉSULTATS GÉNÉRAUX 126
5.3.1 PARTICIPATION 127
5.3.2 CONSTANCE DE LA ROUTINE 128
5.3.3 ÉVÉNEMENTS DE VIE 131
5.3.4 LIMITES DE L'ÉTUDE 135
CONCLUSION 139
BIBLIOGRAPHIE 143
APPENDICE A – CERTIFICATS D'APPROBATION ÉTHIQUE 149
A.1 APPROBATION ÉTHIQUE 2019-035 149
A.2 APPROBATION ÉTHIQUE 2019-035 (RENOUVELLEMENT) 153
A.3 APPROBATION ÉTHIQUE UQAC 2019-035 156
A.4 APPROBATION ÉTHIQUE UQAC 2019-035 (RENOUVELLEMENT) . . . 160
A.5 APPROBATION ÉTHIQUE 2019-220 163
A.6 APPROBATION ÉTHIQUE 2019-220 (RENOUVELLEMENT) 167

LISTE DES TABLEAUX

TABLEAU 3.1 : COMPOSANTS ÉLECTRONIQUES DU BRACELET À CENTRALE INERTIELLE 44

TABLEAU 3.2 : COMPOSANTS ÉLECTRONIQUES DU CAPTEUR DE VITESSE . 52

TABLEAU 3.3 : COMPOSANTS ÉLECTRONIQUES DU PIR+ 54

TABLEAU 4.1 : INFORMATIONS CONTENUES DANS LE FICHIER DE LOG FORMATÉ 66

TABLEAU 4.2 : CARACTÉRISTIQUES EXTRAITES LORS DU PROCESSUS D'EXTRACTION DE L'INFORMATION. 79

TABLEAU 4.3 : CARACTÉRISTIQUES SPÉCIFIQUEMENT EXTRAITES POUR LES ALGORITHMES DE CHANGEMENT COMPORTEMENTAL 80

TABLEAU 5.1 : PARTICIPATION DES PARTICIPANTS À LA CUEILLETTE 128

TABLEAU 5.2 : ANALYSE DES DIFFÉRENCES ENTRE LA MOYENNE ET LES SCORES DE CHANGEMENT COMBINÉS POUR LA SEMAINE 4 ET LA SEMAINE 9 DU PARTICIPANT 2002 133

LISTE DES FIGURES

FIGURE 1.1 – L'HISTORIQUE DES TEMPÉRATURES MOYENNES DE MONTRÉAL POUR L'ANNÉE 2020 6

FIGURE 1.2 – L'HISTORIQUE DE LA VALEUR DE L'ACTION DE LA BANQUE ROYALE DU CANADA POUR L'ANNÉE 2020 6

FIGURE 1.3 – DÉTECTION DE CHANGEMENT SUR L'HISTORIQUE DE LA VALEUR DE L'ACTION DE LA BANQUE ROYALE DU CANADA POUR L'ANNÉE 2020 8

FIGURE 2.1 – REPRÉSENTATION DE D POUR $A = 3$ ET $M = 24$ 26

FIGURE 2.2 – COURBES D'ACTIVITÉS D'UNE JOURNÉE SÉPARÉE EN 288 INTERVALLES D'UNE DURÉE DE 5 MINUTES. 27

FIGURE 3.1 – COMPOSANTS ÉLECTRONIQUES DU BRACELET 45

FIGURE 3.2 – SCHÉMA DE CONNEXION DU BRACELET 45

FIGURE 3.3 – LE BRACELET À CENTRALE INERTIELLE COMPLET. 46

FIGURE 3.4 – TRANSITIONS ENTRE LES MODES DE FONCTIONNEMENT DU BRACELET À CENTRALE INERTIELLE 48

FIGURE 3.5 – LE CAPTEUR DE VITESSE COMPLET ET INSTALLÉ. 52

FIGURE 3.6 – SCHÉMA DE CONNEXION DU CAPTEUR DE VITESSE 52

FIGURE 3.7 – LE PIR+ COMPLET ET SES CÔNES 55

FIGURE 3.8 – SCHÉMA DE CONNEXION DU PIR+ 55

FIGURE 4.1 – EXTRAIT D'UN FICHIER DE DONNÉES BRUTES 62

FIGURE 4.2 – EXTRAIT D'UN FICHIER DE DONNÉES TRAITÉES 65

FIGURE 4.3 – EXTRAIT D'UN FICHIER D'INFORMATION SUR LES ENTRAÎNEMENTS 77

FIGURE 4.4 – SYNTHÈSE DES TRAITEMENTS EFFECTUÉS SUR LES DONNÉES. 81

FIGURE 5.1 – SYNTHÈSE DE LA MÉTHODE DE SÉLECTION DES FENÊTRES POUR COMPARAISON 92

FIGURE 5.2 – VECTEUR DE PROBABILITÉS CRÉÉ POUR LA MÉTHODE RULSIF-VECTEUR. 97

FIGURE 5.3 – MATRICE DE PROBABILITÉS CRÉÉE POUR LA MÉTHODE RULSIF-MATRICE 98

FIGURE 5.4 – MATRICE DE PROBABILITÉS CRÉÉE POUR LA MÉTHODE RULSIF-CARACTÉRISTIQUES 99

FIGURE 5.5 – EXEMPLE D'UN FICHIER DE SORTIE GÉNÉRÉ À LA SUITE DE LA GÉNÉRATION DES SCORES DE CHANGEMENT 101

FIGURE 5.6 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 481 104

FIGURE 5.7 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME SW-PCAR 105

FIGURE 5.8 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE VECTEUR. 106

FIGURE 5.9 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE 106

FIGURE 5.10 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE DE CARACTÉRISTIQUES. 107

FIGURE 5.11 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 907 109

FIGURE 5.12 – APERÇU DU TEMPS ÉCOULÉ LORSQUE LE BRACELET ÉTAIT HORS DE PORTÉE DU DOMICILE POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 907 . . . 109

FIGURE 5.13 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME SW-PCAR 110

FIGURE 5.14 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE VECTEUR. 111

FIGURE 5.15 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE 111

FIGURE 5.16 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE DE CARACTÉRISTIQUES. 112

FIGURE 5.17 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 1453 . . 113

FIGURE 5.18 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 1692 . . 114

FIGURE 5.19 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME SW-PCAR 116

FIGURE 5.20 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE VECTEUR. 116

FIGURE 5.21 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE 117

FIGURE 5.22 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE DE CARACTÉRISTIQUES. 117

FIGURE 5.23 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 2002 . . 119

FIGURE 5.24 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME SW-PCAR 120

FIGURE 5.25 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE VECTEUR. 121

FIGURE 5.26 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE 121

FIGURE 5.27 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE DE CARACTÉRISTIQUES. 122

FIGURE 5.28 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 2182 . . 123

FIGURE 5.29 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME SW-PCAR 124

FIGURE 5.30 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE VECTEUR. 125

FIGURE 5.31 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE 125

FIGURE 5.32 – GRAPHIQUE ILLUSTRANT LES SCORES DE CHANGEMENT OBTENUS À L'AIDE DE L'ALGORITHME RULSIF EN MODE MATRICE DE CARACTÉRISTIQUES. 126

FIGURE 5.33 – ÉVALUATION DE LA CONSTANCE DE LA ROUTINE POUR CHAQUE PARTICIPANT. 130

FIGURE 5.34 – APERÇU DU TEMPS ÉCOULÉ POUR CHAQUE CATÉGORIE D'ACTION EFFECTUÉE PAR LE PARTICIPANT NUMÉRO 2002 INCLUANT LES ÉVÉNEMENTS DE VIE 132

FIGURE 5.35 – ANALYSE DES DIFFÉRENCES ENTRE LES SCORES DE CHANGEMENT COMBINÉS ET LA MOYENNE POUR LA SEMAINE 4 DU PARTICIPANT 2002 134

FIGURE 5.36 – ANALYSE DES DIFFÉRENCES ENTRE LES SCORES DE CHANGEMENT COMBINÉS ET LA MOYENNE POUR LA SEMAINE 9 DU PARTICIPANT 2002 134

LISTE DES ABRÉVIATIONS

UQAC Université du Québec à Chicoutimi

LIARA Laboratoire d'Intelligence Ambiante pour la Reconnaissance d'Activités

DM1 Dystrophie myotonique de type 1

CPD *Change Point Detection*

RulSIF *Relative unconstrained Least-Square Importance Fitting*

PCB *Printed Circuit Board*

BLE *Bluetooth Low Energy*

PCAR *Permutation-based Change detection in Activity Routine*

DTW *Dynamic Time Warping*

RTOS *Real-Time Operating System*

PIR+ Détecteur de mouvement infrarouge passif à valeur ajoutée

TUG *Timed Up and Go*

MQTT *Message Queuing Telemetry Transport*

DÉDICACE

À mes parents et à mon frère,

J'ai la chance d'appartenir à une famille unie et soudée. Merci d'être là pour moi et de m'encourager dans les projets que j'entreprends. Merci de m'épauler dans les moments plus difficiles et de me donner la motivation nécessaire pour continuer ce que j'entreprends. Sachez que j'en suis très reconnaissant.

REMERCIEMENTS

Je tiens à remercier tous ceux qui m'ont soutenue pendant la réalisation de ma maîtrise et pendant la rédaction de ce mémoire.

Je tiens à remercier toute l'équipe du Laboratoire d'Intelligence Ambiante pour la Reconnaissance d'Activités (LIARA) pour leur soutien et pour m'avoir donné l'opportunité de concrétiser mon projet de maîtrise. Plus particulièrement, j'aimerais remercier mon directeur de recherche Sébastien Gaboury et mon co-directeur de recherche Kevin Bouchard pour leur accompagnement et pour leur expertise précieuse apportée tout au long de ma maîtrise.

Je tiens aussi à remercier le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG) et les Fonds québécois de la recherche sur la nature et les technologies (FRQNT) pour leurs soutien financier pendant la réalisation de mon projet de recherche.

Je remercie également mes collègues du LIARA pour leurs connaissances apportées et pour leur contribution au climat d'entraide et à l'ambiance chaleureuse du laboratoire durant ces deux dernières années.

Finalement, j'aimerais remercier ma famille pour leurs encouragements tout au long de mes études.

CHAPITRE I

INTRODUCTION

1.1 CONTEXTE

La Dystrophie myotonique de type 1 (DM1) est une maladie neuromusculaire incurable dont la plus forte prévalence est au Québec, plus précisément dans la région du Saguenay-Lac-Saint-Jean (Bouchard *et al.*, 1989). Les individus atteints présentent des pertes de force musculaire conséquentes au fil des années. La réalisation d'un programme d'entraînement semble une stratégie intéressante pour contrer les déficiences musculaires chez cette population et il a été démontré comme étant sécuritaire (Gagnon *et al.*, 2008) et pouvant mener à des gains en force maximale lorsque supervisé (Tollback *et al.*, 1999). Malheureusement, à cause de différentes contraintes (ex. : transport, coûts, disponibilités), les programmes d'entraînement sont rarement effectués dans des endroits contrôlés tels des cliniques ou des centres d'entraînement. C'est d'ailleurs pour cette raison que l'entraînement à domicile semble une avenue intéressante pour cette population. Une étude a démontré que des stratégies d'entraînement impliquant une équipe de recherche ou de cliniciens favorisent l'adhésion des patients atteints de DM1 (Aldehag *et al.*, 2013). Assurer la supervision des programmes d'entraînement à domicile afin de maximiser l'adhésion des patients comporte de nombreux défis. Les visites à domicile et les appels téléphoniques à intervalles réguliers sont coûteux en plus de ne pas être optimaux. Le soutien par une technologie d'assistance des personnes atteintes de DM1 peut être une solution efficace pour les accompagner et les motiver dans

la réalisation de leur programme d'entraînement à domicile, à condition de développer une solution peu coûteuse qui guide les patients efficacement dans leur programme. C'est d'ailleurs un sujet étudié au Laboratoire d'Intelligence Ambiante pour la Reconnaissance d'Activités de l'Université du Québec à Chicoutimi dans le cadre du projet Acti-DM1.

Dans le cadre de ce projet, un prototype de technologie d'assistance à domicile pour les personnes atteintes de DM1 a été conçu et récemment testé en contexte réel sur une période de 10 semaines. La technologie développée consiste en différents dispositifs à coût faible permettant de guider les patients dans la réalisation de leur programme d'entraînement à domicile tout en recueillant des informations permettant d'effectuer le suivi de leur progression. Un bracelet à centrale inertielle est porté par le participant et reconnaît les séquences d'entraînement prévues au programme ainsi que certaines activités physiques effectuées hors du programme (ex. : marcher, courir, etc.). Une unité centrale de traitement connectée à un petit haut-parleur effectue la liaison avec le bracelet et rappelle les exercices quotidiens à effectuer. Un article rédigé par notre équipe expliquant la technologie est actuellement en seconde révision dans une revue avec arbitrage par les pairs (Chapron *et al.*, 2021b).

En plus d'avoir guidé les patients dans leur programme d'entraînement d'une durée de 10 semaines, la technologie a recueilli des données d'activité physique quotidiennement à l'aide du bracelet à centrale inertielle (ex. : activités physiques effectuées, sorties du domicile, etc.) Les travaux présentés dans ce mémoire portent sur l'analyse de ces données afin de détecter automatiquement des changements de comportement associés à l'amélioration de la capacité physique ou à un événement de vie (ex. : chute, chirurgie). Pour chaque participant effectuant

le programme d'entraînement à domicile assisté par la technologie, il sera question de traiter les données recueillies pour extraire l'information pertinente afin d'effectuer l'analyse des changements comportementaux à l'aide de plusieurs algorithmes reconnus.

1.2 PROBLÉMATIQUE

La présente section est dédiée à la présentation du problème de recherche. Dans un premier temps, les concepts clés seront définis. Ceux-ci vont permettre une compréhension plus pointue du domaine de recherche traité dans ce mémoire. Ensuite, les différentes hypothèses de recherche seront spécifiées et justifiées. Finalement, l'énoncé de la question de recherche sera présenté.

1.2.1 TERMINOLOGIE ET CONCEPTS

Cette recherche est intimement liée à l'étude des changements comportementaux ainsi qu'à la DM1. Plusieurs concepts clés seront donc définis dans cette section.

1.2.1.1 DYSTROPHIE MYOTONIQUE DE TYPE 1

Selon le portail d'information sur la maladie de Steinert (DM1) de l'association APMIS, la DM1 est définie comme suit :

> La dystrophie myotonique de type1 (DM1), appelée aussi maladie de Steinert ou myopathie de Steinert, est une maladie génétique neuromusculaire rare hé-

> réditaire, transmise génétiquement par un parent porteur. [...] En général elle s'aggrave à chaque génération (apparition de plus en plus précoce, symptômes plus nombreux et plus importants). [...] Comme toutes les myopathies, la maladie de Steinert se caractérise par des atteintes des muscles (affaiblissement : « dystrophie », troubles du tonus : « myotonie »), mais aussi par des dérèglements d'autres organes (appareil cardiorespiratoire, système digestif, yeux, système nerveux, système hormonal . . .). [...] Elle se présente sous différentes « formes » (forme asymptomatique/tardive, forme adulte commune, forme infantile, forme néonatale/congénitale), bien que l'anomalie génétique à l'origine de la maladie soit la même. L'évolution de la maladie chez un individu particulier n'est pas prévisible, et l'espérance de vie est normale s'il y a une bonne prise en charge des symptômes (APIMS, 2020).

Cette maladie représente la myopathie la plus fréquente chez l'adulte. Le muscle squelettique est particulièrement affecté (Orngreen *et al.*, 2005), (Tollback *et al.*, 1999) et les individus atteints présentent des pertes de force musculaire maximale évaluées entre 24.5 à 52.8% sur une période de 9 ans (Gagnon *et al.*, 2018). De plus, elle connait sa plus forte prévalence au Québec, plus précisément dans la région du Saguenay-Lac-Saint-Jean (Bouchard *et al.*, 1989).

1.2.1.2 ANALYSE DES CHANGEMENTS COMPORTEMENTAUX

Pour bien comprendre ce qu'est l'analyse des changements comportementaux, il est d'abord nécessaire de définir ce qu'est une série temporelle ainsi que la détection de rupture.

Série temporelle

Une série temporelle est composée d'une suite de données observées à des points spécifiques dans le temps. Elle prend la forme d'une suite de valeurs ordonnées chronologiquement représentant l'évolution d'une mesure au fil du temps. Plusieurs approches existent pour analyser les séries temporelles afin de répondre à certaines questions (ex. : prédire les tendances futures, analyser le comportement des données dans le temps, etc.). La discipline liée à cette analyse s'appelle l'analyse de séries temporelles (Shumway & Stoffer, 2017).

De nos jours, on retrouve des séries temporelles dans la plupart des domaines. Afin d'illustrer le concept, quelques exemples de série temporelle seront présentés.

Le premier exemple se rapporte à la météorologie. Une série temporelle simple pourrait être générée par un thermomètre qui enregistre la température toutes les heures. La mesure ici est la température en degré Celsius et celle-ci évolue dans le temps. Avec le traitement de cette série temporelle, il est possible d'extraire, par exemple, la température moyenne par jour et de l'illustrer sous la forme d'un graphique sur une période de plusieurs jours (voir Figure 1.1).

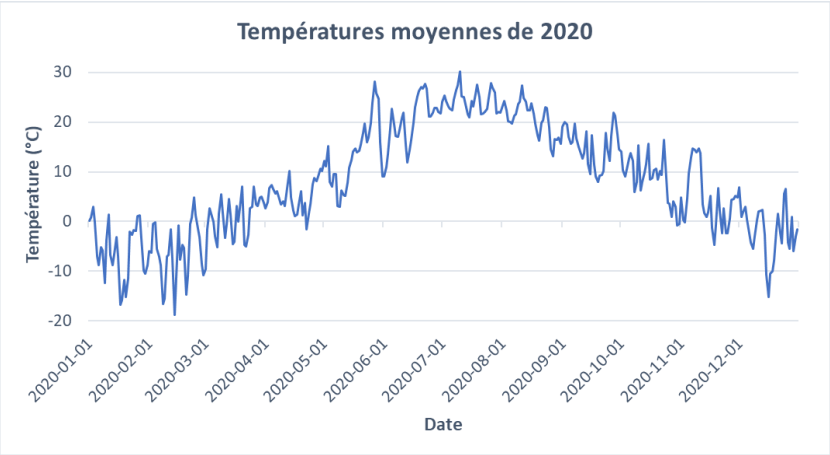


Figure 1.1 : L'historique des températures moyennes de Montréal pour l'année 2020

Le deuxième exemple se rapporte à la bourse. Prenons le prix d'une action : celui-ci varie dans le temps, nous avons donc affaire à une série temporelle (voir Figure 1.2).

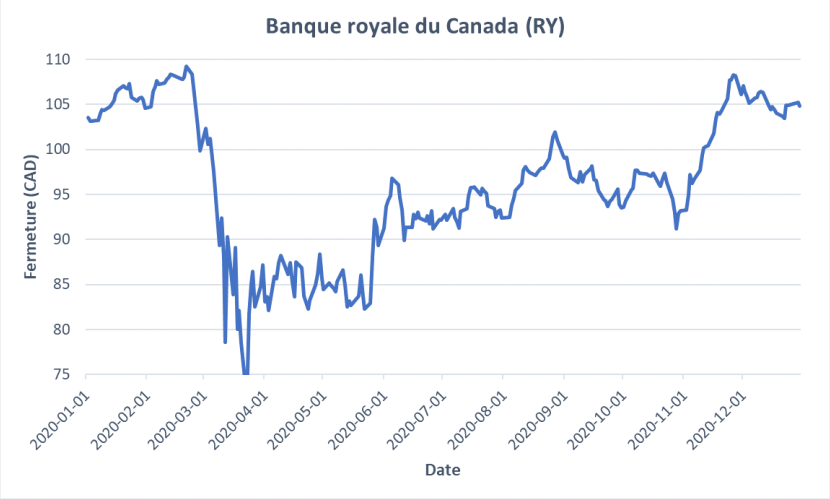


Figure 1.2 : L'historique de la valeur de l'action de la Banque Royale du Canada pour l'année 2020

Détection de ruptures

La détection de ruptures, ou *Change Point Detection* (CPD) en anglais, est l'art de détecter les changements abrupts dans une série temporelle (Aminikhanghahi & Cook, 2017). Le but est de détecter s'il y a un changement et à quel moment le changement se produit si c'est le cas. Les critères pour détecter un changement dépendent du problème traité. Par exemple, la méthode pourrait se baser sur la tendance observée dans la série temporelle et reconnaître un changement à la suite d'une variation de celle-ci. On trouve plusieurs domaines d'application à cette discipline, voici quelques exemples.

— L'analyse des changements climatiques : l'analyse du climat et l'élaboration de modèles de prédiction qui utilisent le CPD est devenue très populaire ces dernières années (Reeves *et al.*, 2007).

— Suivi de la condition médicale : des méthodes de CPD sont étudiées en lien avec certains contextes médicaux tels le sommeil (Staudacher *et al.*, 2005), l'épilepsie (Malladi *et al.*, 2013) ainsi que l'analyse de l'activité cérébrale (Aminikhanghahi & Cook, 2017).

— Analyse des activités d'humains : l'analyse des données de capteurs présents dans une maison intelligente afin de reconnaître une transition d'activité ou encore pour détecter les changements comportementaux de l'habitant peuvent se classer comme étant des applications de CPD (Shumway & Stoffer, 2017).

L'aspect théorique de la détection de ruptures n'est pas le sujet de ce mémoire, le but ici est plutôt d'introduire le concept. Reprenons le graphique sur l'évolution du cours de l'action

de la Banque Royale du Canada présenté à la Figure 1.2. Visuellement, il est possible de voir des changements abrupts dans cette série temporelle. Avec la détection de ruptures, deux points de changements pourraient être détectés automatiquement, ceux-ci sont indiqués dans la Figure 1.3. Cette détection peut être effectuée en utilisant diverses méthodes mathématiques ou algorithmiques. Certaines de ces méthodes seront décrites dans le chapitre 2 de ce mémoire.

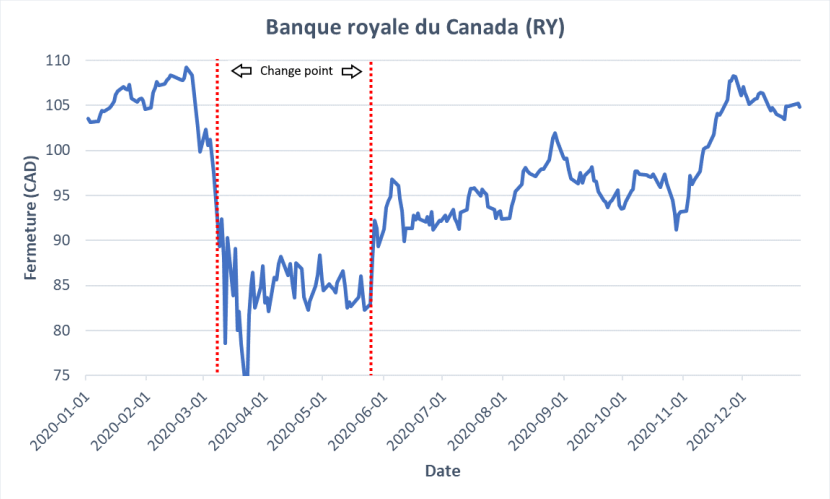


Figure 1.3 : Détection de changement sur l'historique de la valeur de l'action de la Banque Royale du Canada pour l'année 2020

Analyse des changements comportementaux

L'analyse des changements comportementaux peut être vue comme une forme de CPD reliée à l'analyse des activités d'une personne où les données de capteurs constituent des séries temporelles à la suite d'un certain prétraitement. Dans ce cas, un changement abrupt sur les données pourrait être expliqué par un événement de vie.

Événement de vie

Un événement de vie se rapporte à un événement qui se déroule à un moment précis et qui est assez important pour causer un changement comportemental pour un individu. Par exemple, un événement de vie pourrait être l'apparition d'une maladie grave ou encore la naissance d'un enfant. Ces événements sont généralement assez importants pour induire un changement dans la routine d'un individu, ce qui peut les rendre détectables au niveau des données lorsque l'analyse des changements comportementaux est effectuée.

1.2.2 HYPOTHÈSES

Le sujet de recherche implique plusieurs variables qui étaient incertaines au tout début du projet. C'est pourquoi nous avons défini et limité le champ de recherche en posant les hypothèses suivantes.

Hypothèse H1

H1 : **Au moins la moitié des participants à la cueillette de données y participeront de manière rigoureuse et constante.**

La cueillette de données est un facteur clé du projet de recherche. Nous savons que la cueillette de données est un défi important en recherche et que la participation n'est pas toujours parfaite. C'est pourquoi, dès le début, nous avions posé une hypothèse pessimiste de 50% de participation assidue, ce qui laisse une marge de manoeuvre importante pour le reste

du projet quant à la quantité et à la qualité des données espérées.

Hypothèse H2

H2 : **Il y aura un changement de performance musculaire entre le début et la fin du programme d'entraînement.**

Comme il a été démontré que la condition physique des patients atteints de DM1 bénéficie généralement d'un programme d'entraînement adapté et supervisée (Tollback *et al.*, 1999), nous posons l'hypothèse qu'un changement de conditions physiques sera observé entre le début et le fin du programme d'entraînement pour au moins un participant.

Hypothèse H3

H3 : **Au moins un événement de vie notable aura lieu pendant la cueillette de données.**

Comme la cueillette de données s'est déroulée sur une période d'un peu moins de 3 mois avec dix participants atteints de DM1, les chances qu'un événement de vie se produise sont plutôt importantes. En effet, plus la population est grande et plus la durée de la cueillette est élevée, plus les chances qu'un événement de vie se produise augmentent. De plus, la clientèle participant à cette cueillette est un peu plus à risque de subir un événement de santé. Par exemple, les personnes atteintes de DM1 sont plus à risque d'effectuer une chute ou de développer de nouveaux symptômes tels des problèmes cardiaques ou l'apnée du sommeil

(Johnson, 2019). Nous avons donc posé l'hypothèse qu'au moins un événement de vie aura lieu lors de la cueillette de données.

1.2.3 QUESTION DE RECHERCHE

L'objectif de ce projet de recherche est d'implémenter et d'appliquer les algorithmes de détection automatique de changements les plus largement rencontrés dans des situations se rapprochant le plus au contexte du projet Acti-DM1, c'est-à-dire, à une clientèle atteinte de la DM1 qui effectue des entraînements à domicile afin d'améliorer leur condition physique. La question de recherche qui sera traitée sera donc la suivante :

« Comment détecter efficacement un changement comportemental causé par un événement de vie ou l'amélioration de la condition physique suite aux exercices à domicile pour des personnes atteintes de DM1 ? ».

1.3 CONTRIBUTION

Afin de répondre à la question de recherche posée dans la section précédente, le projet de recherche contribue à la discipline de détection de changement de trois façons différentes.

Dans un premier temps, une cueillette de données en situation réelle sera effectuée. Ces données seront traitées et anonymisées afin de les rendre utilisables pour l'exécution d'algorithmes de changement comportemental. Ceux-ci seront d'ailleurs rendus disponibles à la communauté, ce qui constitue une contribution importante dans le domaine.

Ensuite, les algorithmes de détection de changement implémentés au cours de ce projet ont été appliqués dans un nouveau contexte. En effet, c'est la première fois qu'on applique ces algorithmes à une clientèle souffrant de DM1. Il sera donc intéressant pour la communauté scientifique d'en apprendre plus sur les résultats obtenus et les enjeux spécifiques à ce contexte.

Finalement, les travaux présentés dans ce mémoire comportent une contribution directe au projet Acti-DM1. En effet, l'implémentation des algorithmes les plus récents et populaires dans la littérature pour la détection de changement comportemental se greffe à la technologie novatrice d'assistance à domicile développée dans la cadre du projet Acti-DM1. Ces algorithmes pourront donc être réutilisés lors de nouvelles études en lien avec ce projet.

1.4 MÉTHODOLOGIE

Le projet de recherche est séparé en trois phases échelonnées sur une période de 24 mois. Cette section résume l'objectif de chacune de ces phases.

La première phase a tout d'abord été l'acquisition des connaissances du domaine de recherche. Pour ce faire, une revue de littérature sur les méthodes de détection de changement comportemental et de leur contexte d'application a été effectuée. Cela a permis d'extraire l'information clé, tels les algorithmes constituant l'état de l'art pour ce type de problème dans le domaine ainsi que les contextes pour lesquels ces algorithmes ont été appliqués. C'est à la suite de cette étape qu'il a été possible de situer le contexte du projet de recherche et de choisir quels algorithmes allaient être implémentés.

La deuxième phase a ensuite consisté à préparer et à effectuer la cueillette des données. Pour commencer, la planification de la cueillette avec les différents intervenants ainsi que l'élaboration de la demande de certification éthique pour le projet ont été effectuées. Mon rôle a surtout été la préparation logistique. Plus précisément, le développement d'une version miniaturisée du bracelet à centrale inertielle pour utilisation lors de la cueillette a été effectué. De plus, la réplication des différents capteurs (assemblage, soudure des composants) a été faite. Finalement, lors de la phase de cueillette des données, j'ai participé au suivi et aux réparations de matériel lorsque cela était nécessaire.

La troisième phase a finalement été d'exploiter les données recueillies lors de la phase 2. Il a donc été question de nettoyer et de formater les données brutes pour ensuite extraire l'information utile. Ces informations ont permis l'extraction des caractéristiques qui ont servi d'entrées aux algorithmes de détection de changement comportemental. D'ailleurs, ces algorithmes ont été implémentés et exécutés lors de la phase 3, ce qui a créé un jeu de résultats qui a été analysé et discuté.

1.5 ORGANISATION DU DOCUMENT

Ce mémoire est organisé comme suit. Le présent chapitre avait pour but de présenter le projet de recherche portant sur l'analyse des changements comportementaux dans le cadre du projet Acti-DM1. Le deuxième chapitre est composé de la revue de littérature des travaux scientifiques pertinents. Les principaux algorithmes de détection de changement comportemental y seront décrits ainsi que leurs contextes d'applications principaux. Le troisième chapitre a

pour but de présenter la cueillette de données. Plus précisément, les informations sur la mise en place et le déroulement de ladite cueillette seront incluses ainsi que les informations sur les différents dispositifs matériels utilisés. Le quatrième chapitre va décrire les données recueillies lors de la cueillette ainsi que les défis que représente leur traitement. Il sera aussi question de présenter en détail les traitements effectués sur les données afin d'extraire l'information utile en vue d'exécuter les algorithmes de détection de changement comportemental. Le cinquième chapitre a pour but de présenter les méthodes de détection de changement comportemental, les résultats obtenus ainsi que l'analyse et interprétation de ces résultats. Finalement, le sixième chapitre présentera la conclusion générale de ce mémoire et proposera quelques pistes à explorer comme travaux futurs.

CHAPITRE II

ÉTAT DE L'ART

Le second chapitre de ce mémoire a pour but d'explorer les différents travaux scientifiques en lien avec le sujet de recherche proposé. Dans un premier temps, il sera question de présenter les principales méthodes que l'on retrouve dans la littérature. Ensuite, les algorithmes tels que RulSIF, PCAR et sw-PCAR seront détaillés. Finalement, certains travaux récents qui mettent à l'oeuvre ces algorithmes seront présentés. Il sera donc possible de voir dans quels contextes chacun de ces algorithmes a été utilisé ainsi que les résultats obtenus.

2.1 ALGORITHMES

La revue de littérature a commencé avec une recherche générale sur les algorithmes de détection de changement et sur leur contexte d'application. Différentes recherches d'articles scientifiques ont été effectuées sous différents moteurs de recherches. La plupart du temps, il est possible de retrouver le texte complet des articles à l'aide du service de la bibliothèque de l'Université du Québec à Chicoutimi (UQAC). Lorsque le portrait général des différentes méthodes et de leurs contextes d'application a été effectué, seulement celles qui pouvaient êtres appliquées dans ce projet et qui ont montrées une bonne efficacité dans la détection de changements comportementaux ont été retenues. Les algorithmes RulSIF, PCAR et sw-PCAR ont donc été explorés et utilisés dans le cadre de ce projet. D'ailleurs, ces algorithmes seront expliqués dans ce chapitre. Mentionnons tout d'abord les principales méthodes de détection

de changements qui ont été rencontrées lors de l'élaboration de la revue de littérature.

Il existe deux grandes catégories de méthodes pour la détection de changements : les méthodes supervisées et les méthodes non supervisées.

2.1.1 MÉTHODES SUPERVISÉES

Les méthodes supervisées consistent à entraîner des classificateurs à détecter un changement entre les états d'une série temporelle. Afin que ces méthodes soient efficaces, un grand nombre de données diversifiées doivent être utilisées comme exemples pour l'entraînement.

Lorsque le nombre d'états est connu, il est possible d'utiliser un classificateur multi-classe afin de détecter le début et la fin de chaque classe dans la série temporelle (Cleland *et al.*, 2014). Souvent, des fenêtres glissantes sont utilisées sur les données de la série temporelle afin de couvrir les points de changements possibles (Ortiz Laguna *et al.*, 2011). On dénombre plusieurs classificateurs multi-classes tels les arbres de décisions (Salzberg, 1994; Kim *et al.*, 2005; Breiman, 2001), les machines à vecteurs de support (Steinwart & Christmann, 2008; Reddy *et al.*, 2010), la classification naïve bayésienne (Cheng & Greiner, 2013), la méthode des voisins les plus proches (Gao & Wang, 2007; Wei & Keogh, 2006), le modèle de Markov caché (Rabiner & Juang, 1986) et bien d'autres.

On peut aussi traiter la détection de changement en tant que problème de classification binaire. La première classe représente alors tous les points de changements entre deux états et la seconde représente un point intra état où il n'y a pas de changement. L'utilisation de ce type

de classification peut s'avérer complexe par le fait qu'un très grand nombre de transitions est possible et que les données vont souffrir d'un fort déséquilibre vu le nombre de transitions beaucoup moins élevées que le nombre de points intra états. Les méthodes telles les machines à vecteurs de support ou la classification naïve bayésienne ont été appliquées pour la détection de changements en utilisant la classification binaire (Feuz *et al.*, 2015).

Virtual Classifier est une autre méthode supervisée qui peut non seulement détecter un changement, mais aussi interpréter la nature de ce changement (Hido *et al.*, 2008). Un identifiant quelconque (ex. : -1) est associé aux données de la première fenêtre et un autre identifiant (ex. : +1) aux données de la deuxième fenêtre. Un classificateur (ex. : arbre de décision) est alors entraîné pour tenter de distinguer ces deux fenêtres. Si le classificateur parvient à les distinguer correctement, cela implique qu'une différence existe entre ces deux fenêtres et l'explication de la différence peut être retrouvée selon le modèle entraîné.

Les méthodes supervisées doivent être entraînées avec un bon nombre de données et, sauf pour l'algorithme *Virtual Classifier*, les points de changement doivent être connus à l'avance pour les données d'entraînement. C'est pourquoi elles ne peuvent être appliquées à l'étude présentée dans ce mémoire et n'ont donc pas été retenues. En effet, le nombre de données n'est pas assez grand et les points de changements ne sont pas connus à l'avance.

2.1.2 MÉTHODES NON SUPERVISÉES

Les méthodes non supervisées permettent de trouver des points de changement dans une série temporelle même si les données ne sont pas identifiées préalablement (ex. : les points de changements ne sont pas connus à l'avance). Elles se basent généralement sur des caractéristiques statistiques obtenues à partir des données de la série temporelle (Aminikhanghahi & Cook, 2017).

Les méthodes basées sur le test du rapport de vraisemblance (*likelihood-ratio* en anglais) sont des méthodes qui formulent le problème de détection de changements en problème lié au domaine des probabilités et statistiques. Le principe réside dans l'hypothèse que, si les densités de probabilité de deux échantillons sont essentiellement différentes, c'est qu'un changement est observé. On peut distinguer ces méthodes en deux sous-catégories : les méthodes calculant la densité de probabilité de chaque échantillon pour ensuite calculer leur ratio et les méthodes calculant le ratio de densité directement. Les méthodes calculant le ratio de densité directement sont plus simples que les méthodes calculant les deux densités de probabilité. De ce fait, les méthodes récentes utilisent généralement ce principe (Aminikhanghahi & Cook, 2017). Parmi celles-ci, on peut mentionner les algorithmes KLIEP (Sugiyama *et al.*, 2008), ulSIF (Kanamori *et al.*, 2009), RulSIF (Yamada *et al.*, 2013) et SEP (Aminikhanghahi *et al.*, 2019). D'ailleurs, RulSIF a été exploré plus en détail et utilisé lors de l'analyse des changements comportementaux (voir Section 2.1.3).

D'autres méthodes comme *Subspace Identification* (SI) (Kawahara *et al.*, 2007) re-

présentent une série temporelle en espace d'états et tentent de détecter un changement en identifiant les distances de cet espace d'état (Aminikhanghahi & Cook, 2017).

On retrouve aussi des méthodes probabilistes comme l'algorithme bayésien de Adams & MacKay (2007) qui exploite la probabilité d'un changement dans la série temporelle à l'endroit courant selon le temps écoulé depuis le dernier changement.

Certains algorithmes utilisent le partitionnement de données (*clustering* en anglais) afin de détecter un changement. Ces algorithmes tentent de partitionner la série temporelle de manière à séparer les différents états se trouvant dans celle-ci et ainsi retrouver les points de changements avec l'analyse de ces partitions. Parmi les méthodes exploitant ce principe, mentionnons MDL (Rakthanmanon *et al.*, 2011) et Shapelet (Zakaria *et al.*, 2012).

Afin de procéder au choix des algorithmes à utiliser pour ce projet de recherche, les articles récents effectuant l'évaluation de ceux-ci dans un contexte similaire à celui traité dans ce mémoire ont été analysés. Premièrement, comme mentionnés plus tôt, seulement les algorithmes non supervisés seront retenus. La Section 2.2 présente les différentes études retenues qui mettent en collaboration certains algorithmes de détection de changement comportementaux dans un contexte similaire. D'ailleurs, ce sont ces articles qui ont guidé le choix des algorithmes utilisés pour le projet de recherche présenté dans ce mémoire : RulSIF, PCAR et sw-PCAR ont été retenus.

2.1.3 RULSIF

L'algorithme *Relative unconstrained Least-Square Importance Fitting* (RulSIF) (Yamada *et al.*, 2013) permet de comparer deux distributions de probabilités afin d'en tirer une mesure de dissimilarité. En effet, celle-ci estime directement les ratios de densité des deux distributions de probabilités spécifiées en entrée, ce qui permet ensuite d'estimer la mesure de dissimilarité de Pearson (Pearson, 1900). La mesure de dissimilarité résultante va alors nous permettre d'effectuer l'analyse des changements comportementaux.

Bien que la théorie mathématique sur les distributions statistiques est hors de propos dans ce mémoire, il sera tout de même question de présenter le fonctionnement général de cette méthode.

Supposons que nous avons des échantillons $\{x_i\}_{i=1}^{n}$ d'une distribution de probabilités de d-dimensions P avec une densité $p(x)$ et des échantillons $\{x_j\}_{j=1}^{n'}$ d'une distribution de probabilités de d-dimensions Q avec une densité $q(x)$. Le but est de comparer ces deux distributions sous-jacentes P et Q à partir seulement des deux ensembles d'échantillons $\{x_i\}_{i=1}^{n}$ et $\{x_j\}_{j=1}^{n'}$.

On définit le ratio de densité α-relatif de $p(x)$ et $q(x)$ par la relation suivante :

$$r_\alpha(x) = \frac{p(x)}{\alpha p(x) + (1-\alpha)q(x)}, \qquad 0 \leq \alpha < 1. \tag{2.1}$$

Le coefficient de dissimilarité relatif de Pearson entre les densités de probabilité $p(x)$ et $q(x)$ se calcule via l'équation suivante :

$$PE_\alpha[p(x),q(x)] = \frac{1}{2}\int \left(\frac{p(x)}{\alpha p(x)+(1-\alpha)q(x)}\right)^2 \cdot (\alpha p(x)+(1-\alpha)q(x))\,dx. \tag{2.2}$$

Si $PE_\alpha[p(x),q(x)] = 0$ alors les densités de probabilités sont égales.

L'idée est de modéliser le vrai ratio de densité α-relatif, noté $r_\alpha(x)$ par un modèle de la forme :

$$r(x;\theta) = \sum_{l=1}^{n} \theta_l\, K(x,x_l), \tag{2.3}$$

où $\theta := (\theta_1,\ldots,\theta_n)$ sont des paramètres qui seront déterminés (appris) automatiquement à partir des échantillons. La fonction $K(x,x')$ est une fonction noyau définie par :

$$K(x,x') = \exp\left(\frac{-||x-x'||^2}{2\sigma^2}\right), \quad \sigma > 0, \tag{2.4}$$

et $||\cdot||$ désigne la distance euclidienne.

Les paramètres θ dans le modèle $r(x;\theta)$ sont déterminés de manière à ce que l'erreur quadratique J soit minimisée :

$$J(\theta) = \frac{\alpha}{2n}\sum_{i=1}^{n} r^2(x_i;\theta) + \frac{(1-\alpha)}{2n'}\sum_{i=1}^{n'} r^2(x'_i;\theta) - \frac{1}{n}\sum_{i=1}^{n} r(x_i;\theta). \tag{2.5}$$

D'un point de vue de l'algèbre matricielle, on peut formuler le problème ainsi. La fonction $r(x;\theta)$ peut s'écrire de la manière suivante :

$$r(x;\theta) = \sum_{l=1}^{n} \theta_l\, K(x,x_l) = \theta^T k(x) \tag{2.6}$$

où A^T dénote la transposée de la matrice A, $\theta = [\theta_1,\ldots,\theta_n]^T$ et $k(x) = [K(x,x_1),\ldots,K(x,x_n)]^T$.

La fonction de coût peut se formuler comme suit :

$$J(\theta) = \frac{1}{2}\theta^T \hat{H}\theta - \theta^T \hat{h} \tag{2.7}$$

où

$$\hat{H} = \frac{\alpha}{n}\sum_{i=1}^{n} k(x_i)k(x_i)^T + \frac{(1-\alpha)}{n'}\sum_{i=1}^{n'} k(x'_i)k(x'_i)^T \tag{2.8}$$

et

$$\hat{h} = \frac{1}{n}\sum_{i=1}^{n} k(x_i). \tag{2.9}$$

Ainsi, on cherche à résoudre le problème d'optimisation suivant :

$$\hat{\theta} = \underset{\theta\in\mathbb{R}^n}{\arg\min}\left[\frac{1}{2}\theta^T \hat{H}\theta - \theta^T \hat{h} + \frac{\lambda}{2}\theta^T\theta\right] \tag{2.10}$$

où $\lambda > 0$ est un terme de pénalisation (régularisation).

La solution analytique de ce problème d'optimisation (Yamada *et al.*, 2013) est donnée par la relation suivante :

$$\hat{\theta} = \left(\hat{H} + \lambda I_n\right)^{-1}\hat{h} \tag{2.11}$$

où A^{-1} désigne la matrice inverse de A et I_n représente la matrice identité de dimension $n \times n$.

Par conséquent, en remplaçant θ de l'équation (2.6) par $\hat{\theta}$, on obtient l'estimation du ratio de densité α-relatif :

$$\hat{r}_\alpha(x) = \hat{\theta}^T k(x). \tag{2.12}$$

Finalement, l'estimation de la mesure de dissimilarité de Pearson relative est obtenue en évaluant l'expression suivante :

$$\widehat{PE}_\alpha = \frac{-\alpha}{2n}\sum_{i=1}^{n}\hat{r}_\alpha(x_i)^2 - \frac{(1-\alpha)}{2n'}\sum_{i=1}^{n'}\hat{r}_\alpha(x'_i)^2 + \frac{1}{n}\sum_{i=1}^{n}\hat{r}_\alpha(x_i) - \frac{1}{2}. \tag{2.13}$$

On peut retrouver une implémentation de la méthode RulSIF en code source libre (langage Python) [1] ainsi qu'un implémentation en R [2].

Les méthodes d'estimation du ratio de densité (ex. : ulSIF) sont généralement sans limites, ce qui peut causer un ratio de densité infini. RulSIF a été créé spécifiquement pour pallier à ce problème en introduisant le paramètre α (Yamada *et al.*, 2013) qui permet de limiter les valeurs du ratio afin d'éviter ce problème.

De plus, un élément très intéressant avec RulSIF c'est qu'il permet de gérer les échantillons de données à plusieurs dimensions. Cette particularité rend possible la comparaison de plusieurs formes de données différentes. Par exemple, il est possible d'évaluer la mesure de dissimilarité entre deux vecteurs ou entre deux matrices. Cet avantage sera d'ailleurs exploité lors de l'application de cet algorithme au projet de recherche présenté dans ce mémoire.

2.1.4 PCAR

Permutation-based Change detection in Activity Routine (PCAR) est une méthode qui utilise la modélisation généralisée de la routine journalière d'une personne selon les données des capteurs environnants afin de détecter des changements dans celles-ci (Dawadi *et al.*,

1. https://github.com/hoxo-m/densratio_py
2. https://github.com/hoxo-m/densratio

2016). Cette méthode comporte donc deux fonctions principales, soit la modélisation de la routine et la détection de changement dans cette routine.

Courbes d'activités

Afin de modéliser la routine journalière d'un individu, la méthode se base sur des données préalablement traitées par un algorithme de reconnaissance d'activités de manière à les transformer en suite d'activités (ex. : dormir, manger, sortir du domicile, etc.) ordonnées dans le temps. Ces données sont ensuite utilisées afin de modéliser les courbes d'activités journalières représentant la routine de l'individu. Plusieurs activités ont une certaine prévalence pour un moment de la journée. Par exemple, l'activité "dormir" est normalement très présente la nuit et l'activité "déjeuner" est très présente dans le début de l'avant-midi. Les courbes d'activités exploitent ce phénomène. En effet, une journée est séparée en m périodes de durée égale, ce qui permet de calculer la distribution de probabilité sur les activités pour chaque intervalle de la journée. La probabilité associée à une activité dans un intervalle de temps correspond au nombre de minutes consacrées à réaliser cette activité divisée par la durée en minutes de la période. Enfin, l'assemblage des distributions de probabilité de chaque intervalle pour une journée nous donne les courbes d'activité de cette journée.

Notons que la routine peut varier d'une journée à l'autre. Par exemple, la personne ne prend pas nécessairement ses repas exactement aux mêmes heures et ne va pas au lit tous les jours au même moment. Les courbes d'activité doivent donc prendre en compte ces petites variations qui se produisent d'une journée à l'autre. Pour ce faire, les courbes de N journées

sont agrégées pour former une courbe d'activité agrégée qui généralise mieux la routine. L'agrégation des journées consiste à effectuer la moyenne des N journées contenues dans l'intervalle afin de créer une journée agrégée qui est représentative de cet intervalle. Pour ce faire, la moyenne des probabilités d'activité contenue dans chaque intervalle m est effectuée entre les N journées de l'intervalle. Par exemple, si chaque journée comporte 24 périodes (m = 24) et nous voulons agréger sur 30 jours ($N = 30$), la moyenne des périodes est calculée afin de transformer la fenêtre originale de taille 30 par 24 en fenêtre agrégée de taille 24 dont chaque élément contient les probabilités moyennes calculées à partir des probabilités contenues dans les 30 intervalles pour cette position.

Nous avons posé comme m étant le nombre d'intervalles dans une journée et N comme étant le nombre de journées de la fenêtre d'agrégation. Posons a comme étant le nombre d'activités possibles, D comme étant l'ensemble des distributions de probabilité d'une journée et $\hat{D}$ comme étant l'agrégation de D sur N journées. Dans ce cas, D comporte m éléments et chaque élément de D comporte a éléments. En d'autres mots, chaque élément de D est une distribution de probabilité pour un intervalle de temps contenu dans une journée. La Figure 2.1 illustre bien le rôle des différentes variables.

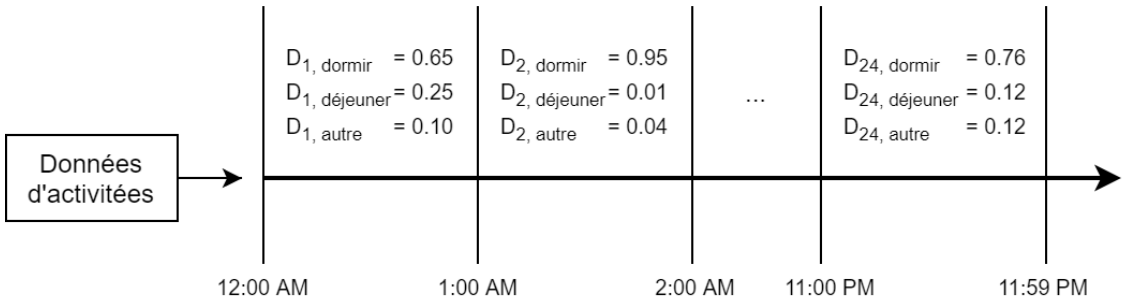


Figure 2.1 : Représentation de D pour $a = 3$ et $m = 24$

L'un des avantages importants des courbes d'activité agrégée réside par le fait qu'elles permettent facilement de visualiser la routine généralisée d'un individu. En effet, si l'on modélise la probabilité d'occurrence de chaque activité pour chaque intervalle de la journée, le modèle démontre les courbes de façon explicite (voir Figure 2.2).

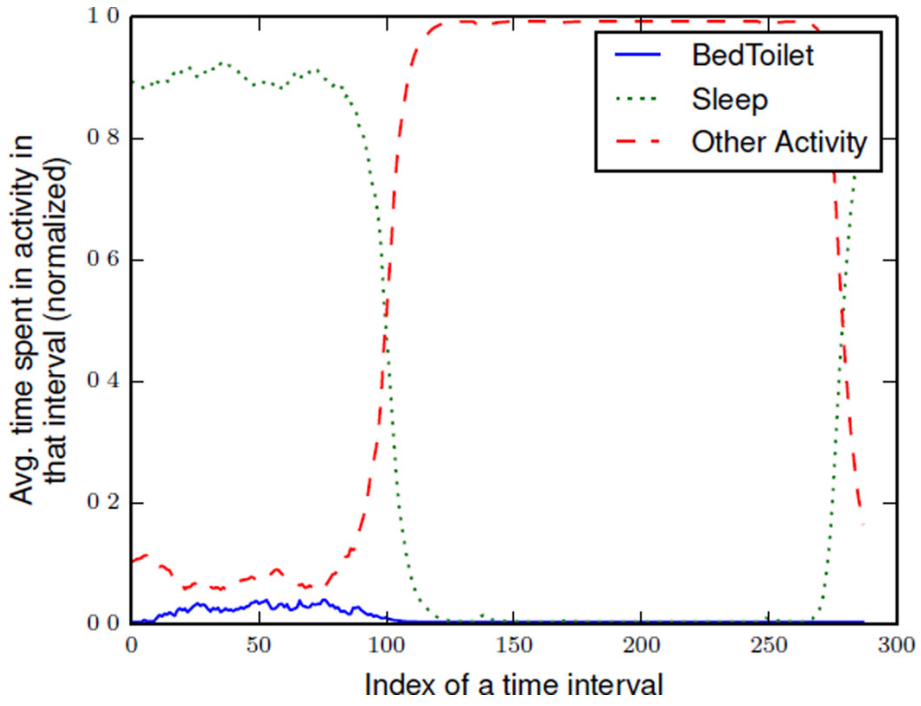


Figure 2.2 : Courbes d'activités d'une journée séparée en 288 intervalles d'une durée de 5 minutes (Dawadi *et al.*, 2016)

Détection de changements

La détection de changement s'effectue en comparant deux fenêtres agrégées $\hat{D}$ appartenant au même jeu de données. Par exemple, si nous avons des données d'habitat intelligent d'un résident pour un an, la première fenêtre pourrait représenter les six premiers mois et la deuxième fenêtre les six derniers mois. Dans ce cas, il serait possible de comparer les deux fenêtres entre elles afin de déterminer si un changement conséquent survient entre les deux routines généralisées. Cependant, pour que la comparaison soit effectuée, il faut d'abord aligner les distributions des deux fenêtres agrégées. Deux méthodes sont disponibles pour effectuer cela :

— Alignement basé sur les périodes : cette méthode est en fait la plus simple, car aucune action supplémentaire n'est requise pour aligner les deux fenêtres. En effet, chaque intervalle de temps est comparé avec la correspondante de l'autre fenêtre agrégée. Par exemple, la distribution d'activité pour la période no. 1 de la première fenêtre est comparée à celle de la période no. 1 de la deuxième fenêtre et ainsi de suite.

— Alignement dynamique basé sur la déformation du temps : cette méthode permet d'accommoder les changements mineurs entre les deux courbes d'activité agrégées. Ces changements mineurs peuvent, par exemple, correspondre au décalage d'au plus une heure du début ou de la fin d'une activité ou bien à l'augmentation ou la diminution d'au plus une heure de la durée d'une activité. L'algorithme utilisé par cette méthode est *Dynamic Time Warping* (DTW) qu'il est possible de voir plus en détail dans l'article original (Dawadi *et al.*, 2016).

Une fois les fenêtres alignées correctement, il est maintenant temps d'exécuter la routine de détection de changement. PCAR utilise un test de permutation à deux échantillons (Efron & Tibshirani, 1994) afin de détecter si un changement important a lieu entre deux ensembles de distribution (i.e., les distributions d'activités provenant des fenêtres agrégées). Voici donc les grandes étapes effectuées par l'algorithme PCAR pour la détection de changement :

1. Calcul de la mesure de divergence symétrique de Kullback–Leibler entre les deux distributions d'activité agrégées originales. La mesure de Kullback–Leibler est tirée des statistiques et a pour objectif de calculer la distance entre deux distributions de

probabilités (Kullback & Leibler, 1951). Pour PCAR, la version symétrique est utilisée et est définie comme suit : $SD_{KL}(D_1||D_2) = D_{KL}(D_1||D_2) + D_{KL}(D_2||D_1)$.

2. Obtient la distribution empirique de la mesure statistique (divergence symétrique de Kullback–Leibler). Pour ce faire, il y a tout d'abord un mélange des distributions d'activité entre les deux fenêtres agrégées. Ensuite, un recalcul de la mesure statistique est effectué. Ces deux étapes sont réalisées en boucle jusqu'à ce que la distribution empirique soit complète.

3. Détermine si un changement important existe entre les deux fenêtres agrégées en comptant le nombre de fois que les mesures statistiques de la distribution empirique sont plus grandes ou égales à l'originale. Lorsque le compte est relativement bas, PCAR considère qu'un changement important est présent entre les deux fenêtres agrégées.

Finalement, l'algorithme 2.1 résume le fonctionnement de PCAR tel que décrit précédemment.

Données : $C1, C1$, deux collections de courbes d'activité

Résultat : CS, le score de changement

$CA1 = AgregerCourbe(C1)$

$CA2 = AgregerCourbe(C2)$

$empDist = DistributionEmpirique(CA1, CA2)$

$AL = AlignerCourbes(CA1, CA2)$

pour chaque *pairs d'alignement* (i, j) **dans** AL **faire**

    Calculer $divKL = SD_{kl}(D_{1,i}, D_{2,j})$ entre les distributions d'activités $D_{1,i} \in CA1$ et $D_{2,j} \in CA2$;

    Déterminer si le changement est significatif selon $divKL$ et la distribution empirique $empDist$ (comptage);

    si *le changement est significatif* **alors**

        $CS = CS + 1$;

    fin

fin

retourner CS

Algorithme 2.1 : Fonctionnement de l'algorithme PCAR

Les explications précédentes sont bien sûr un résumé de la méthode afin d'en expliquer le fonctionnement général. Pour plus de détails sur son fonctionnement, veuillez vous référer à l'article de Dawadi *et al.* (2016).

2.1.5 SW-PCAR

L'algorithme sw-PCAR est une modification de PCAR qui permet d'utiliser de plus petites fenêtres de temps. En effet, PCAR fonctionne correctement lorsqu'il est utilisé sur des données s'échelonnant sur une longue période, mais devient plus limité, voire inapplicable, pour des périodes plus courtes (Sprint *et al.*, 2016b). De plus, sw-PCAR a été adapté afin de comparer des nombres scalaires (ex. : données d'activité physique, nombre de pas).

L'algorithme sw-PCAR est basé sur le même principe de permutation que PCAR, mais son fonctionnement est différent :

1. Dans un premier temps, comme pour PCAR, il est nécessaire de choisir deux fenêtres de données qui seront utilisées pour la comparaison. Une moyenne est effectuée sur chacune des fenêtres, ce qui donne deux fenêtres agrégées. Par exemple, si nous avons 24 périodes dans une journée et la fenêtre originale comporte 30 jours, une moyenne pour chaque période est effectuée pour transformer la fenêtre originale de taille 30 par 24 en fenêtre agrégée de taille 24 dont chaque élément correspond à la moyenne des 30 intervalles pour cette position.
2. La mesure de distance symétrique de Kullback–Leibler est ensuite appliquée entre les deux fenêtres agrégées, ce qui donne le score de changement noté *CS*.
3. Les deux fenêtres agrégées sont concaténées pour ensuite être mélangées de façon aléatoire. Une fois cela fait, la fenêtre concaténée est séparée en deux, ce qui va donner deux nouvelles fenêtres de tailles égales pour lesquelles le score de changement est

calculé puis stocké. Cette étape est répétée *N* fois afin de produire un ensemble de scores de changement exhaustif.

4. La dernière étape consiste à effectuer un test de signification statistique à partir du score de changement original et de l'ensemble créé à l'étape précédente. La méthode utilisée est tirée du *boxplot-based outlier detection* (Maimon & Rokach, 2005). Dans un premier temps, l'ensemble des scores de changement obtenus à l'étape 3 est trié en ordre croissant pour y extraire le premier et troisième quartile. Ensuite, si le score de changement original est plus grand que le troisième quartile + 1.5 fois l'étendue interquartile, le changement est jugé conséquent.

L'algorithme 2.2 résume le fonctionnement de sw-PCAR tel que décrit précédemment.

Données : $W1, W2$, deux fenêtres de données ; N, nombre de permutations

Résultat : CS, le score de changement ; Sig, Indique si CS est significatif

$WA1 = AgregerFenetre(W1)$

$WA2 = AgregerFenetre(W2)$

$CS = SD_{kl}(WA1, WA2)$

$V =$ vecteur de taille N

$k = 0$

tant que $k < N$ **faire**

    Mélanger les intervalles de temps de WA1 et WA2;

    Générer deux nouvelles fenêtres WMA1 et WMA2 à partir des intervalles mélangées;

    $SDKL = SD_{kl}(WMA1, WMA2)$;

    Ajouter la valeur de $SDKL$ au vecteur V;

    $k = k + 1$;

fin

$sig = BoxplotOutlierDetection(CS, V)$

retourner CS, Sig

Algorithme 2.2 : Fonctionnement de l'algorithme sw-PCAR

Pour plus de détails, le lecteur est invité à se référer à l'article de Sprint *et al.* (2016b).

2.2 CONTEXTES D'APPLICATION

La section précédente présente les algorithmes récents trouvés dans la littérature scientifique s'apprêtant au domaine de recherche étudié. Dans cette section, il sera question de décrire les différents travaux qui reprennent ces algorithmes dans un contexte similaire à celui étudié dans ce mémoire.

2.2.1 HABITATS INTELLIGENTS

Les habitats intelligents sont généralement source d'un grand nombre d'informations nous renseignant sur les résidents qui y habitent. Généralement, les informations temporelles d'activité de la vie quotidienne sont les plus communes. Des recherches ont justement exploité ces informations riches afin de détecter des changements dans la routine de l'habitant (Dawadi *et al.*, 2016), ce qui peut d'ailleurs être causé par un événement de vie important (ex. : chirurgie, changement familial, etc.) (Sprint *et al.*, 2016a).

Une équipe de chercheurs a recueilli des données de 18 maisons intelligentes où résidaient des personnes âgées sur une durée de deux ans (Dawadi *et al.*, 2016). L'algorithme PCAR a été utilisé afin de construire les courbes d'activités représentatives de la routine de chacun des 18 individus. Parallèlement, 2 fois par année, chacun des participants était évalué à l'aide des tests TUG (Podsiadlo & Richardson, 1991) et RBANS (Randolph, 1998). Ces tests avaient comme objectifs d'évaluer et de quantifier leurs habilitées physiques et cognitives tout le long de la cueillette afin de suivre l'évolution de la santé globale de chacun. Le test

Timed Up and Go (TUG) consiste à mesurer le temps requis pour se lever de position assise, parcourir trois mètres et retourner s'asseoir alors que RBANS est un ensemble de tests ayant comme objectif l'évaluation de la santé cognitive d'un individu. L'objectif était de corréler les changements perçus au niveau des tests d'évaluation à ceux détectés automatiquement en effectuant la détection de changement sur les routines à l'aide de PCAR. Les résultats indiquent qu'il y a bel et bien une corrélation entre les changements observés à l'aide des tests et les changements détectés à l'aide de l'algorithme exploitant les données issues d'une multitude de capteurs installés dans le domicile, ce qui démontre que certains changements au niveau de la capacité physique ou cognitive peuvent affecter la routine d'un individu. Il a aussi été remarqué que les scores de changement étaient plus élevés après la première année, surtout chez les individus en moins bonne santé, ce qui donne un autre indice quant à la corrélation entre la santé et la routine d'une personne.

Une autre équipe de chercheurs a utilisé les données issues de maisons intelligentes afin de comparer la routine de personnes souffrant de problème cognitif (ex. : démence) à la routine de personnes en bonne santé (Sprint *et al.*, 2021). Chaque groupe comportait 7 personnes et les facteurs tels l'âge et l'éducation étaient semblables entre ces deux groupes. Les 7 personnes du premier groupe souffraient de problèmes cognitifs tels la démence, et les 7 personnes du deuxième groupe étaient en bonne santé. La cueillette de données s'est effectuée sur une durée d'un mois. Au terme de la cueillette, les algorithmes *Virtual Classifier*, PCAR et SEP ont été appliquées pour déceler les différences entre la routine des deux groupes. Les résultats indiquent des différences plus marquées pour la vitesse de marche, pour l'activité

« prendre son bain » et pour l'activité « effectuer une sortie du domicile ».

D'autres chercheurs ont tiré avantage de données issues d'habitats intelligents afin de mettre à l'épreuve certains algorithmes de détection de changement comportemental vis-à-vis des événements de vie importants observés chez des personnes plus âgées (Sprint *et al.*, 2016a). L'étude de cas comportait des données d'habitats intelligents pour trois personnes pour lesquels un évènement de vie s'est produit (traitement de radiothérapie, insomnie et chute). Les algorithmes utilisés lors de cette recherche sont sw-PCAR, RulSIF et *Virtual Classifier*. Pour la suite, nous allons nous concentrer seulement sur les algorithmes qui nous intéressent, soit RulSIF et sw-PCAR.

1. La première personne a suivi un traitement de radiothérapie suivant un diagnostic de cancer du poumon pendant le troisième mois de collecte de données. Les algorithmes de détection de changement sw-PCAR et RulSIF ont bel et bien détecté des changements importants suivant le début du traitement. On dénote d'ailleurs une modification des habitudes pour les activités « Dormir » et « Sortir du domicile, » ce qui est vraisemblablement lié au traitement de radiothérapie.
2. La deuxième personne a reçu un diagnostic d'insomnie durant la cueillette de données. Les deux algorithmes ont efficacement détecté des changements importants dans la routine de la participante suivant le diagnostic. Les changements les plus importants ont été notés au niveau du sommeil et des sorties du domicile qui ont tous deux diminué.
3. La troisième personne a effectué une chute. Les effets de cette chute sur la routine

détectée par l'habitat intelligent n'a pas été conséquent selon les résultats des algorithmes sw-PCAR et RulSIF. Cela démontre que la routine doit être affectée de manière plus conséquente afin qu'un changement soit détecté.

Cette section démontre clairement l'intérêt d'utiliser des algorithmes de changements comportementaux dans le contexte des habitats intelligents. Les recherches pertinentes retenues pour illustrer l'utilité des algorithmes de détection de changement comportemental dans ce contexte démontrent des résultats encourageants effectués sur des données bien réelles. Ceux-ci peuvent être appliqués soit pour quantifier les changements de routine qui s'imposent à long terme avec le vieillissement ou l'apparition de problèmes de santé, soit pour détecter un événement de vie important qui affecte la routine de manière conséquente.

2.2.2 DÉTECTION DE CHANGEMENT AU NIVEAU DE L'ACTIVITÉ PHYSIQUE

Avec la popularité grandissante des dispositifs portables tels les montres intelligentes, le monitorage de l'activité physique est de plus en plus accessible. En effet, la plupart de ces dispositifs comportent des capteurs pouvant compter les pas effectués, la fréquence cardiaque ou parfois même mesurer la distance parcourue à l'aide de données GPS. Ces données sont très utiles pour dresser un portrait général de l'activité physique effectuée par la personne qui porte le dispositif.

Une équipe de chercheurs s'est d'ailleurs penchée sur la détection automatique des changements au niveau de l'activité physique en exploitant ces données. L'objectif est d'utiliser

les informations recueillies à l'aide des capteurs portatifs que l'on retrouve dans les montres connectées afin de détecter les changements d'activité physique et d'estimer leur cause, ce qui peut aider à la motivation vis-à-vis de l'accomplissement d'un objectif personnel de santé (Sprint *et al.*, 2016b). Une cueillette de données a été effectuée auprès de 11 personnes dans le cadre d'une intervention de santé visant des personnes de 55 ans et plus jugées à risque de développer certaines maladies tels la démence ou certains troubles cardio-vasculaires. Lors de la première et de la dernière semaine de l'intervention, les participants ont porté une montre FitBit Flex [3] qui enregistrait leurs données d'activité physique 24 heures sur 24. Entre ces deux semaines, le bracelet n'était pas porté, mais de l'éducation sur les bonnes habitudes de vie et l'importance de l'exercice a été effectuée de même que l'élaboration d'objectifs personnels de santé. Chaque semaine, les participants effectuaient une auto-évaluation pour quantifier le niveau d'atteinte des objectifs fixés. Finalement, des algorithmes de détection de changement ont utilisé le comptage des pas effectués comme données de base pour détecter les changements.

Les algorithmes *Virtual Classifier*, RulSIF et sw-PCAR ont été utilisés, mais la présente section va se concentrer sur les résultats obtenus avec les deux derniers. Premièrement, l'expérience démontre bien les enjeux des cueillettes de données en situation réelle. En effet, les données comportent beaucoup de bruit et des différences marquées ont été observées entre l'atteinte des objectifs rapportés via l'auto-évaluation et les résultats réellement observés à l'aide de la technologie. Par exemple, un participant a rapporté avoir totalement atteint son

3. https://www.fitbit.com/au/flex2

objectif de nombre de pas, mais une diminution de ce nombre a été observée entre la première et la dernière semaine. De plus, certaines données sont manquantes en raison du non-port de la montre pour certaines journées, ce qui introduit une difficulté supplémentaire quant à l'exploitation des données réelles.

Malgré ces difficultés, il a été possible de remarquer plusieurs changements au niveau de l'activité physique pour la plupart des personnes. L'algorithme sw-PCAR a détecté des changements pour cinq participants lorsqu'il a été utilisé avec des périodes d'une durée de 5 minutes, pour quatre participants avec des périodes de 10 et 15 minutes et pour trois participants avec des périodes entre 20 et 40 minutes. L'algorithme RulSIF a détecté un peu moins de changements avec une seule détection pour plusieurs tailles de périodes différentes. Le plus grand nombre de détections se retrouve pour des périodes dont la durée est un peu plus élevée avec 3 pour une période d'une durée de 35 minutes et 4 pour une période d'une durée de 45 minutes. Somme toute, sw-PCAR à détecté un total de 42 changements et RulSIF un total de 22 changements.

Une autre étude a exploité les données d'activités physiques recueillies à l'aide d'un bracelet FitBit [4] afin d'explorer l'évolution journalière de certains patients en cours de réhabilitation (Sprint *et al.*, 2017). Deux groupes différents ont été formés : un premier groupe de 9 personnes en bonne santé vivant normalement et un deuxième groupe de 15 personnes en cours de réhabilitation. L'objectif du groupe de contrôle était de former un groupe pour lequel aucun changement significatif n'était prévu mis à part les variabilités quotidiennes afin de le

4. https://www.fitbit.com/global/fr-ca/home

comparer au groupe de patients en cours de réhabilitation, et ainsi observer les différences quant au niveau de changement. Les algorithmes tels que sw-PCAR et *Virtual Classifier* ont été appliqués sur les données de chaque participant afin de déceler et d'évaluer les changements au niveau de l'activité physique (nombre de pas, battements de coeur, périodes de repos, etc.) observée pour la période traitée (2 semaines pour le groupe de contrôle et de 4 à 25 jours pour les participants du groupe en réhabilitation). Les résultats n'ont démontré aucun changement significatif pour les personnes dans le groupe de contrôle alors que plusieurs changements ont été détectés pour plusieurs personnes en cours de réhabilitation. Pour ce groupe, une grande variabilité est observée vis-à-vis les différentes périodes comparées et les changements les plus importants se retrouvent vers la fin du traitement, principalement au niveau du nombre de pas effectués.

Avec ces exemples concrets d'application des algorithmes de détection de changement comportemental dans le contexte de l'activité physique, il a été possible de remarquer la pertinence des méthodes choisies ainsi que le degré d'efficacité de chacune selon la taille de l'intervalle de temps impliqué.

2.3 DISCUSSION

Bien que plusieurs techniques aient été développées au fil des années pour la détection de changement, les algorithmes sw-PCAR, PCAR et RulSIF démontrent une efficacité qui a été récemment prouvée dans des contextes d'études en lien avec le sujet étudié. Dans la cadre de ce mémoire, RulSIF et sw-PCAR seront utilisés pour suivre l'évolution des caractéristiques

reliées à l'activité physique des participants de la cueillette de données. Comme mentionné plus tôt, PCAR n'est pas conçu pour traiter de courtes périodes, c'est pourquoi il ne peut pas être appliqué à cette étude. Cependant, sa version améliorée (sw-PCAR) est tout à fait adaptée. De plus, RulSIF nous permettra d'évaluer les changements de plusieurs manières différentes selon le type d'entrée, ce qui va augmenter l'expressivité des résultats.

CHAPITRE III

CUEILLETTE DE DONNÉES

La phase 2 du projet de recherche décrit dans ce mémoire consistait à recueillir des données en situation réelle chez des individus atteints de DM1. Cette cueillette de données avait pour but de tester la solution technologique d'assistance à domicile développée au LIARA et de recueillir des données essentielles pour la suite de ce projet. Ce chapitre a donc pour objectif de présenter en quoi consistait, plus précisément, la cueillette de données en commençant par introduire les différents dispositifs qui ont été déployés chez les participants pour terminer en expliquant les détails de sa mise en place et de son déroulement.

3.1 DISPOSITIFS

Les dispositifs développés pour effectuer la cueillette de données se greffent à la solution technologique du projet Acti-DM1. Le capteur principal de cette technologie d'assistance est un bracelet à centrale inertielle qui enregistre les données inertielles au besoin et reconnaît l'activité effectuée en temps réel par la personne qui le porte. Ce bracelet transmet ces informations à une unité centrale de traitement qui s'occupe de les recueillir et de gérer le déroulement des exercices à domicile. Finalement, deux autres dispositifs s'ajoutent à cela : un capteur de vitesse de marche et un capteur de mouvement. Cette section est consacrée à la présentation détaillée de chacun de ces dispositifs.

3.1.1 BRACELET À CENTRALE INERTIELLE

Le bracelet à centrale inertielle est un composant important de la solution d'assistance à domicile qui a été conçue et assemblée au LIARA. En effet, une première version avait été développée précédemment (Chapron *et al.*, 2018) et une deuxième version de taille plus petite comprenant certaines fonctionnalités supplémentaires a été conçue pour la cueillette de données du projet Acti-DM1. Cette sous-section va donc se consacrer à la présentation de la version 2 de ce bracelet, en commençant par détailler l'aspect matériel pour ensuite exposer son fonctionnement.

Composition matérielle

Le bracelet est muni d'une centrale inertielle comportant un accéléromètre, un gyroscope et un magnétomètre. Chacune de ces unités retourne des mesures sur trois axes (X, Y et Z) pour un total de 9 signaux. Le bracelet est aussi muni d'un microcontrôleur supportant la connectivité Bluetooth, d'une batterie ayant une capacité de 350 milliampères-heure et d'un contrôleur de charge. Le tout est intégré dans un boîtier imprimé en 3D. Le Tableau 3.1 présente les différents composants électroniques utilisés.

Tableau 3.1 : Composants électroniques du bracelet à centrale inertielle

Nom	Description	Quantité
RedBearLab BLE Nano 2	Microcontrôleur Bluetooth	1
Sparkfun LSM9DS1	Centrale inertielle 9 axes	1
SparkFun Battery Babysitter	Contrôleur de charge	1
Batterie	Batterie 350 mah	1
Bouton	Bouton 6mm SMD	1
3mm DEL	DEL de 3mm verte	1
Résistance	Résistance 220 ohm SMD	2

Afin d'assembler ces différents composants électroniques, un circuit imprimé, ou *Printed Circuit Board* (PCB) en anglais, a été conçu et créé au sein du LIARA. Cela a été rendu possible à l'aide du logiciel Autodesk Eagle [5] développé par la compagnie Autodesk et conçu spécifiquement pour la conception et le design de circuits imprimés. Le PCB développé a finalement été produit à l'aide de l'imprimante à PCB Voltera [6] disponible au laboratoire. La Figure 3.1 montre un prototype de circuit imprimé développé pour le bracelet à centrale inertielle et la Figure 3.2 présente le schéma de connexion.

5. https://www.autodesk.com/products/eagle/

6. https://www.voltera.io/product/pcb-printer

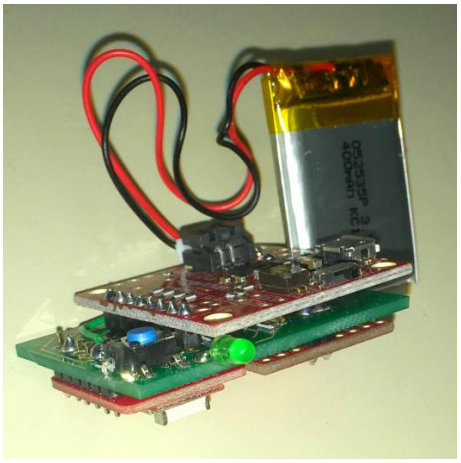


Figure 3.1 : Composants électronique du bracelet (Chapron *et al.*, 2021a)

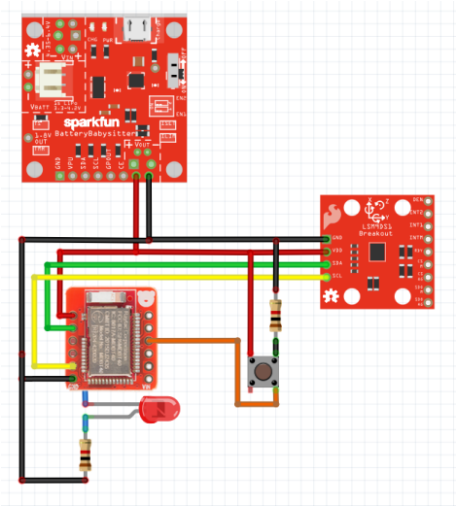


Figure 3.2 : Schéma de connexion du bracelet

Enfin, un boîtier a été créé à l'aide du logiciel de conception 3D Fusion 360 [7] développé par Autodesk. Suivant sa conception, plusieurs exemplaires ont été imprimés en 3D avec l'imprimante Lulzbot Taz 6 [8] disponible au LIARA. La Figure 3.3 montre le bracelet complet assemblé dans le boîtier en question.

7. https://www.autodesk.ca/fr/products/fusion-360/
8. https://www.lulzbot.com/store/printers/lulzbot-taz-6

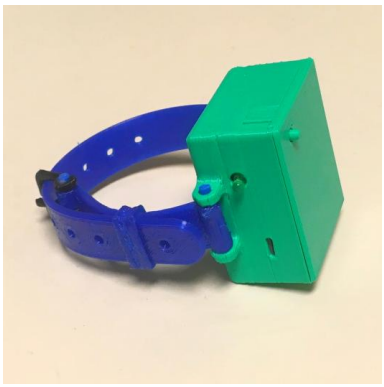


Figure 3.3 : Le bracelet à centrale inertielle complet (Chapron *et al.*, 2021a)

Fonctionnement

Le programme embarqué du bracelet à centrale inertielle a été écrit en langage C++ et est basé sur le système d'exploitation temps réel ou, en anglais, *Real-Time Operating System* (RTOS) Mbed [9] spécialement conçu pour le développement de dispositifs se rapportant à l'internet des objets. La connectivité *Bluetooth Low Energy* (BLE) est assurée par la puce nRF52832 de Nordic Semiconductor [10]. C'est une puce très populaire dans le domaine des dispositifs portables, car elle permet une connectivité sans-fil à l'aide de la technologie BLE avec une consommation énergétique faible. On la retrouve d'ailleurs dans plusieurs produits portables tels que des montres connectées.

Lorsqu'il est porté, ce bracelet permet de recueillir plusieurs données brutes, mais aussi de générer de l'information à partir de ces données à l'aide d'algorithmes légers s'exécutant directement sur le dispositif. Selon le mode de fonctionnement, soit les données brutes, soit

9. https://os.mbed.com/

10. https://www.nordicsemi.com/Products/Low-power-short-range-wireless/nRF52832

l'information prétraitée est envoyée à l'unité centrale de traitement via le lien de communication sans-fil. Voici les informations qui peuvent être transmises à l'unité centrale de traitement :

1. Activité courante : le logiciel embarqué exploite un arbre de décision afin de reconnaître l'activité courante selon les données de la centrale inertielle. Les activités pouvant être reconnues sont les suivantes :
   - Inactif : la personne est inactive
   - Assis-Debout : la personne se lève
   - Debout-Assis : la personne s'assoie
   - Courir : la personne court
   - Marcher : la personne marche
2. Statut de mouvement : le statut de mouvement est reconnu à partir des données provenant de la centrale à inertie et nous indique si oui ou non il y a eu un mouvement effectué, aussi minime soit-il. C'est cet indicateur qui va servir à détecter le port du bracelet par la suite.
3. Indicateur de charge : indique si oui ou non le bracelet est en cours de chargement.
4. Niveau de batterie : indique le niveau de batterie en pourcentage.

Selon le mode de fonctionnement courant du bracelet, ces informations ne sont pas transmises de la même façon à l'unité centrale de traitement. En effet, le fonctionnement du bracelet peut être subdivisé en trois modes distincts :

1. Mode de fonctionnement régulier (en ligne)
2. Mode de fonctionnement déconnecté (hors ligne)
3. Mode entraînement

Le mode de fonctionnement régulier est actif lorsque le bracelet est à portée de l'unité centrale de traitement, c'est-à-dire lorsqu'il se trouve à l'intérieur du domicile du participant. Si le dispositif devient hors de portée de l'unité centrale de traitement (ex. : le participant est hors du domicile), le mode déconnecté devient alors actif. Finalement, après un appui de 5 secondes sur le bouton du bracelet, le mode exercice est déclenché. La Figure 3.4 montre les transitions entre les différents modes.

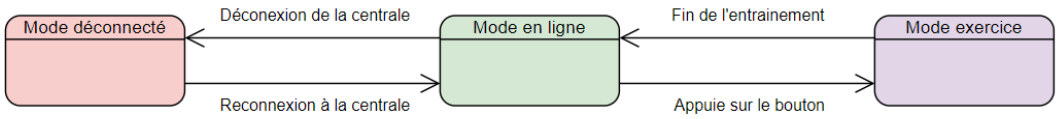


Figure 3.4 : Transitions entre les modes de fonctionnement du bracelet à centrale inertielle

Au cours du mode de fonctionnement régulier, le bracelet est connecté à l'unité centrale de traitement via une liaison BLE, ce qui permet d'y envoyer directement les informations générées par ce dispositif. En effet, les activités reconnues (Inactif, Assis-Debout, Debout-Assis, Marcher, Courir), le statut de mouvement, le statut de charge ainsi que le niveau de charge sont transmis aussitôt qu'ils sont générés.

De plus, lorsque le logiciel embarqué détecte une déconnexion de la liaison BLE avec

l'unité centrale de traitement, le mode hors ligne est activé. Au cours de ce mode, les activités sont tout de même reconnues en temps réel, mais sont stockées dans la mémoire interne et renvoyées à l'unité centrale de traitement en lot aussitôt que la connexion est rétablie. Comme la mémoire interne est très limitée, il est impossible de stocker les données de reconnaissance d'activités séquentiellement. En effet, la mémoire serait vite saturée avec l'écriture d'une nouvelle entrée par seconde nous renseignant sur le type d'activité et sur le moment où elle a été effectuée. C'est pourquoi, en mode hors-ligne, l'information sur les activités reconnues n'est pas stockée séquentiellement et les données temporelles ne sont pas sauvegardées. La solution implémentée consiste à sauvegarder un compteur pour chaque activité possible et d'incrémenter le compteur associé à l'activité courante toutes les secondes. Au terme du mode hors-ligne, la mémoire interne va contenir le nombre de secondes pour lesquels chaque activité a été reconnue pendant la période de déconnexion.

Finalement, il est possible de mettre le bracelet en mode exercice avec l'appui d'un bouton. L'activation de ce mode indique alors le début d'une séance d'entraînement à domicile. Les algorithmes pour reconnaître les séquences d'entraînement et guider le participant dans la réalisation de son programme exigent une puissance de calcul que le bracelet à centrale inertielle ne peut pas satisfaire. C'est pourquoi le fonctionnement normal du bracelet est suspendu et les données inertielles brutes sont envoyées en temps réel à l'unité centrale de traitement pendant toute la durée de l'entraînement. Le déroulement des exercices est alors géré par l'unité centrale de traitement vu sa puissance considérablement plus élevée que celle du bracelet. À la fin de la séance d'entraînement, l'unité centrale de traitement indique au

bracelet que le mode exercice est terminé et qu'il peut alors se remettre en mode normal.

3.1.2 UNITÉ CENTRALE DE TRAITEMENT

L'unité centrale de traitement est l'unité de contrôle principale qui lie tous les dispositifs entre eux. Elle est composée d'un Raspberry Pi 3 B+ [11] relié à un haut-parleur. Le Raspberry Pi est un nano-ordinateur de choix pour ce type d'application, car il est assez petit (85 mm par 56 mm) et relativement puissant pour sa taille (processeur 4 coeurs de 1.4 GHz, 1 Go de mémoire vive). Le système d'exploitation officiel Raspbian [12] y est installé et les différents services nécessaires à la cueillette de données sont activés.

Un programme y est installé et incorpore une interface de contrôle pour les différents paramètres de l'expérimentation, effectue le lien entre les dispositifs, traite les données reçues et les garde en mémoire, contrôle les séquences d'entraînement et sauvegarde les données chaque soir.

De plus, lorsque le mode exercice est activé, l'unité centrale de traitement est responsable du traitement des données inertielles brutes afin de reconnaître les exercices/séquences effectués. Chaque série d'exercices est reconnue et un retour vocal est effectué à l'aide du haut-parleur connecté afin de guider les participants dans la réalisation de leurs séances d'entraînement à domicile.

---

11. https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/
12. https://www.raspberrypi.org/software/

Finalement, chaque soir, les informations de la journée (activités effectuées, déconnexions, reconnexions, séquences d'exercices, indicateur de mouvement, etc.) sont envoyées sur un serveur pour des fins de sauvegarde.

3.1.3 CAPTEUR DE VITESSE

Le capteur de vitesse a été développé au cours d'un projet connexe au LIARA (Chapron *et al.*, 2021a). Ce dispositif non intrusif mesure la vitesse de marche des personnes qui passent devant. Il était considéré comme un atout pour la cueillette de données du projet Acti-DM1, car il permet de recueillir des données supplémentaires qui peuvent s'avérer utiles pour la suite du projet (ex. : suivi de la vitesse de marche).

Composition matérielle

Ce dispositif est composé d'un Raspberry Pi Zero W [13], d'un convertisseur numérique-analogique et de trois capteurs de distance infrarouges. Le nano-ordinateur Raspberry Pi Zero W est encore plus petit que la version 3 B+ avec une dimension de 65 mm par 30 mm, mais reste quand même assez puissant avec un processeur de 1 GHz et 512 Mo de mémoire vive. Un avantage majeur de ce nano-ordinateur réside dans le fait qu'il intègre à la fois la connectivité sans-fil WiFi et BLE. Un boîtier imprimé en 3D a aussi été conçu pour abriter ces composants. Le Tableau 3.2 présente les différents composants de ce dispositif.

13. https://www.raspberrypi.org/products/raspberry-pi-zero-w/

Tableau 3.2 : Composants électroniques du capteur de vitesse

Nom	Description	Quantité
Raspberry Pi Zero W	Nano-ordinateur intégrant BLE/WiFi	1
Capteur de distance infrarouge	Modele GP2Y0A02YK0F	3
Convertisseur numérique-analogique	Modèle ADS1115 de Adafruit	1

Le boîtier a été conçu à l'aide du logiciel Autodesk Fusion 360 et a été imprimé en 3D avec l'imprimante Lulzbot Taz 6 disponible au LIARA. La Figure 3.5 montre le capteur de vitesse assemblé et la Figure 3.6 montre son schéma de connexion.

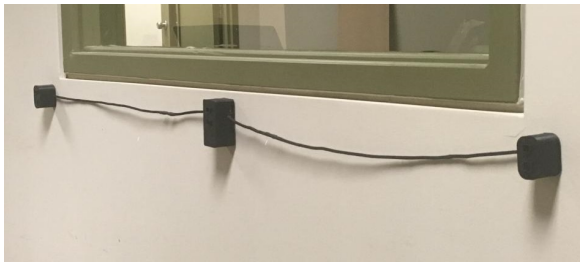


Figure 3.5 : Le capteur de vitesse complet et installé (Chapron *et al.*, 2018)

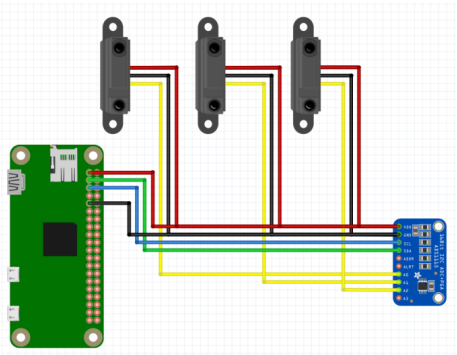


Figure 3.6 : Schéma de connexion du capteur de vitesse (Chapron *et al.*, 2018)

Fonctionnement

Le programme informatique embarqué communique avec trois capteurs de distance infrarouges actifs. Lorsque quelqu'un passe devant ces trois capteurs, un algorithme transforme les lectures effectuées par ceux-ci en vitesse de marche qui est ensuite envoyée à l'unité centrale de traitement à l'aide du canal de communication WiFi. Comme ce dispositif ne constitue pas

une contribution dans le cadre du projet de recherche faisant l'objet de ce mémoire, le lecteur est invité à se référer à l'article de Chapron *et al.* (2021a) pour des détails plus précis sur son fonctionnement.

3.1.4 PIR+

Le détecteur de mouvement infrarouge passif à valeur ajoutée (PIR+) est un dispositif développé dans le cadre de mon projet de maîtrise. La détection de mouvement à l'aide de capteurs infrarouges passifs (PIR) n'est pas quelque chose de nouveau en soi : plusieurs dispositifs utilisent cette technique pour détecter des mouvements. Par contre, il était nécessaire de développer un dispositif se greffant aisément à l'ensemble déjà existant pour la cueillette de données développé au LIARA. De plus, les besoins quant à la détection d'activités en contexte multirésident se faisaient sentir. Ce dispositif a donc intégré une fonction supplémentaire qui facilite son utilisation en contexte multirésident, d'où la notation PIR+. Il devait initialement être inclus dans la cueillette de donnée du projet Acti-DM1. Malheureusement, des contraintes logistiques et pratiques sont apparues lors de la planification de la cueillette et ce dispositif n'a pas été déployé. Il est tout de même intéressant de le présenter, car il comporte une contribution scientifique importante.

Composition matérielle

Tout comme le capteur de vitesse, le PIR+ comporte un nano-ordinateur Raspberry Pi

Zero W qui est connecté à un détecteur de mouvement infrarouge passif Parallax [14] ayant une portée jusqu'à 9 mètres. Les composants électroniques de ce détecteur sont très sensibles aux fortes concentrations d'ondes WiFi, ce qui peut induire de fausses détections. C'est pourquoi un boîtier spécial a été conçu afin de limiter l'exposition du détecteur de mouvement à ces ondes. Le Tableau 3.3 présente les différents composants matériels du dispositif.

Tableau 3.3 : Composants électroniques du PIR+

Nom	Description	Quantité
Raspberry Pi Zero W	Nano-ordinateur intégrant BLE/WiFi	1
Capteur de mouvement	Parallax 555-28027	1
Condensateur	Condensateur de 100µF	1
Résistance	Résistance de 220Ω	1

Le boîtier a été conçu de la même manière que pour le capteur de vitesse et le bracelet, c'est-à-dire à l'aide de Autodesk Fusion 360 et de la Lulzbot Taz 6 du LIARA. Aussi, afin de contrôler l'angle de vision du capteur de mouvement, différents cônes peuvent être installés afin de diminuer l'angle de détection lorsque nécessaire. La Figure 3.7 montre le PIR+ assemblé avec les différents cônes développés et la Figure 3.8 montre le schéma de connexion entre les différents composants électroniques.

14. https://www.parallax.com/product/pir-sensor-with-led-signal/

Figure 3.7 : Le PIR+ complet et ses cônes (Lapointe *et al.*, 2020)

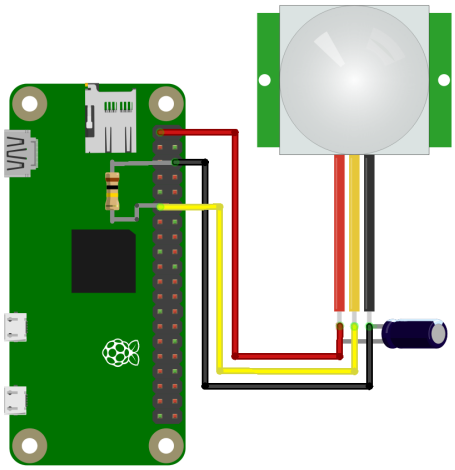


Figure 3.8 : Schéma de connexion du PIR+ (Lapointe *et al.*, 2020)

Fonctionnement

Ce dispositif comporte deux fonctions principales. Dans un premier temps, il permet de capter les mouvements dans son champ de vision, ce qui permet, entre autres, de détecter la présence d'une personne dans l'endroit du domicile concerné. Dans un second temps, lorsqu'il est utilisé de pair avec plus d'un bracelet et avec d'autres capteurs PIR+, il permet d'associer chaque détection de mouvement à la personne qui l'a déclenchée.

Au niveau de la détection de mouvement, le capteur infrarouge passif envoie un signal au nano-ordinateur lorsqu'un mouvement est détecté. Le Raspberry Pi Zero W va alors transmettre un signal indiquant qu'un mouvement a eu lieu à l'unité centrale de traitement via WiFi à l'aide du protocole *Message Queuing Telemetry Transport* (MQTT). Deux types de messages peuvent être envoyés avec cette fonction : un message est envoyé lorsqu'une

détection de mouvement débute et un autre lorsqu'une détection de mouvement se termine.

Au niveau de l'association des détections de mouvement en contexte multirésidents, le fonctionnement se base sur un scan constant des dispositifs émetteurs à portée (ex. : les bracelets). En effet, chaque bracelet émet un identifiant unique via BLE. Ces identifiants ainsi que la force du signal reçu sont captés par chaque dispositif PIR+ installé dans le domicile pour ensuite être formatés et envoyés à l'unité centrale de traitement par WiFi. Finalement, un algorithme s'exécute en temps réel sur l'unité centrale de traitement afin de positionner chaque bracelet relativement à la position de chaque PIR+.

Bien que ce dispositif ait été réalisé lors de ma maîtrise, il ne fait pas l'objet du sujet principal. C'est pourquoi le lecteur est invité à se référer à l'article scientifique décrivant ce dispositif pour plus de détails (Lapointe *et al.*, 2020).

3.2 DÉROULEMENT DE LA CUEILLETTE

Un total de 20 participants divisés en 2 groupes ont pris part à l'expérimentation. Tous étaient atteints de dystrophie myotonique de type 1 et devaient suivre un programme d'entraînement à domicile spécifique. En effet, un programme d'entraînement a été élaboré et devait être effectué pendant toute la durée de la cueillette. Ces entraînements avaient comme objectif de renforcer musculairement les membres inférieurs des patients. Le premier groupe, appelé groupe de contrôle, n'avait pas l'aide de la technologie d'assistance à domicile pour la réalisation de leur programme d'entraînement. Contrairement au premier, le deuxième groupe

prenait part à l'expérimentation avec le soutien technologique développé dans le cadre du projet Acti-DM1. Ce procédé a été appliqué afin de permettre de mesurer l'influence de la technologie vis-à-vis des résultats généraux engendrés par le programme d'entraînement. Il est important de spécifier que cette étude a été approuvée par le comité d'éthique du CIUSSS-SLSJ (Centre Intégré Universitaire de Santé et de Services Sociaux du Saguenay-Lac-St-Jean, Saguenay, Québec, Canada), numéro de dossier 2020-7.

Le déroulement de la cueillette de données s'est effectué en trois temps : la préparation, l'exécution et la finalisation. Les sous-sections suivantes décriront chacune de ces étapes.

3.2.1 PRÉPARATION

Une cueillette de données en situation réelle peut poser un enjeu de planification important. Bien qu'un travail considérable ait été effectué sur le point de la planification et sur l'éthique en recherche par l'équipe responsable du projet Acti-DM1, cette section va plutôt s'occuper de décrire de manière générale les implications directement liées au projet de maîtrise présenté dans ce mémoire.

Mon rôle quant à la préparation de cette cueillette de données a été principalement au niveau du développement informatique. En effet, la première version du bracelet à centrale inertielle qui avait été développé précédemment nécessitait une miniaturisation. Il a donc été question d'une refonte matérielle de celui-ci afin d'inclure des composants plus petits, un nouveau circuit imprimé ainsi qu'un nouveau boîtier imprimé en 3D. De plus, le dispositif

de type PIR+ a été conçu lors de la phase de préparation de cette cueillette de donnée en vue d'être utilisé lors de celle-ci. Finalement, comme chaque dispositif devait être disponible en 10 exemplaires, la réplication du nombre requis a aussi été effectuée lors de cette phase.

Un peu avant le début de la cueillette de données, chaque participant a été évalué deux fois par deux physiothérapeutes différentes au cours de la même semaine. Ces évaluations avaient pour but d'évaluer la force isométrique des membres inférieurs et de recueillir d'autres mesures nécessaires à l'évaluation de la condition musculaire de chaque patient (Lessard *et al.*, 2021).

Aussi, deux semaines avant le début de la cueillette, l'installation et la calibration des dispositifs ont été effectuées par l'équipe responsable du volet technologique. Un bracelet, une unité centrale de traitement et un haut-parleur ont été installés dans chaque résidence et le capteur de vitesse a été installé seulement dans les résidences comportant un endroit opportun (ex. : munie d'un couloir).

3.2.2 EXÉCUTION

Au cours de la durée totale de l'expérimentation de 10 semaines, les participants devaient porter le bracelet toute la journée et le faire charger durant la nuit. De plus, ils devaient suivre le programme d'entraînement prescrit par les physiothérapeutes. D'ailleurs, des rappels étaient effectués par la technologie d'assistance lorsque l'entraînement devait être fait et des communications hebdomadaires avaient lieu entre la physiothérapeute et chaque participant

afin de s'assurer du bon déroulement du programme.

Aussi, chaque nuit, les données de la journée étaient automatiquement envoyées sur un serveur, ce qui nous permettait de valider le bon fonctionnement de la technologie. D'ailleurs, certaines pannes sont survenues au cours de la cueillette et ont été réglées.

Finalement, un journal de bord était tenu par chaque participant. Celui-ci nous renseignait sur plusieurs aspects de la cueillette soit :

- le port du bracelet chaque jour (si oui ou non le bracelet a été porté pour chacune des journées) ;
- la réalisation des exercices (le moment où la séance a été effectuée et le nombre de répétitions pour chaque séquence d'exercices complétée) ;
- les sorties du domicile effectuées (date et heure de la sortie) ;
- les événements de vie qui sont survenus (description de l'événement de vie et dates).

3.2.3 FINALISATION

Au cours de la cueillette de données, certains participants ont dû se retirer pour des raisons personnelles telles que des problèmes de santé, opérations, etc. qui ne sont pas en lien avec l'étude réalisée. Un total de 4 personnes appartenant au groupe de la technologie se sont retirées. Au terme de la durée de la cueillette de 10 semaines, les dispositifs ont été récupérés et les participants restants ont effectués de nouvelles évaluations physiques. En effet, afin d'avoir une mesure comparative entre le début et la fin du programme d'entraînement,

les mêmes physiothérapeutes ont évalué les mêmes mesures liées à la condition musculaire de chaque participant comme il a été décrit dans la sous-section décrivant la préparation. Un article clinique dont je suis un des co-auteurs a aussi été publié au terme de cette étude (Lessard *et al.*, 2021).

CHAPITRE IV

TRAITEMENT DES DONNÉES

Au terme de l'expérimentation de 10 semaines, c'est un total de 22,6 Go de données qui a été recueilli à l'aide des différents dispositifs. Cette quantité massive de données devait tout d'abord être traitée afin de les rendre plus facilement utilisables pour les traitements de détection de changement comportemental, pour les validations et pour l'exportation vers d'autres systèmes, comme, par exemple, l'interface web Acti-DM1. Pour ce faire, elles ont tout d'abord subi un nettoyage et un formatage, ce qui a permis d'éliminer les informations non nécessaires et de les rendre plus facilement lisibles. Une fois cette tâche effectuée, un traitement plus approfondi et plus complexe pouvait être fait, ce qui comprend l'extraction des caractéristiques utiles en vue de les préparer à la détection de changement comportemental. Ce chapitre a donc pour objectif de présenter les données et les différents traitements qui ont été effectués sur celles-ci afin de les préparer pour les algorithmes de détection de changement comportemental.

4.1 CONVERSION DES DONNÉES BRUTES

Les données recueillies au terme de la cueillette apparaissent de manière séquentielle sous forme textuelle et comportent un très grand nombre d'informations. Un traitement de nettoyage et de conversion a donc été effectué en premier lieu afin de rendre ces données beaucoup moins volumineuses et beaucoup plus faciles à lire. Cette section va commencer par

présenter le format brut des données pour ensuite exposer le format standardisé et finalement terminer avec la présentation des étapes qui ont permis d'effectuer cette conversion.

4.1.1 REPRÉSENTATION INITIALE

Les données brutes sont produites par l'unité centrale de traitement de manière continue tous les jours. Un fichier différent par participant est généré chaque jour et chaque ligne est écrite de manière séquentielle. La Figure 4.1 montre un extrait d'un fichier de données brutes.

```
[2019-10-11 00:01:20] (MainHub) 2019/10/11 00:01:20 //DATA queried by E8:44:3E:9D:9B:1D with {"classification":0}.
[2019-10-11 00:01:21] (BLE-DATA) 2019/10/11 00:01:21 [E8:44:3E:9D:9B:1D]: {"immobile_state":255}
[2019-10-11 00:01:21] (MainHub) 2019/10/11 00:01:21 //DATA queried by E8:44:3E:9D:9B:1D with {"immobile_state":255}.
[2019-10-11 00:01:21] (BLE-DATA) 2019/10/11 00:01:21 [E8:44:3E:9D:9B:1D]: {"classification":0}
[2019-10-11 00:01:21] (ACTIVITY-MONITOR) 2019/10/11 00:01:21 Add 0 to activities.
[2019-10-11 00:01:21] (MainHub) 2019/10/11 00:01:21 //DATA queried by E8:44:3E:9D:9B:1D with {"classification":0}.
[2019-10-11 00:01:22] (BLE-DATA) 2019/10/11 00:01:22 [E8:44:3E:9D:9B:1D]: {"immobile_state":255}
[2019-10-11 00:01:22] (BLE-DATA) 2019/10/11 00:01:22 [E8:44:3E:9D:9B:1D]: {"classification":0}
[2019-10-11 00:01:22] (MainHub) 2019/10/11 00:01:22 //DATA queried by E8:44:3E:9D:9B:1D with {"immobile_state":255}.
[2019-10-11 00:01:22] (ACTIVITY-MONITOR) 2019/10/11 00:01:22 Add 0 to activities.
[2019-10-11 00:01:22] (MainHub) 2019/10/11 00:01:22 //DATA queried by E8:44:3E:9D:9B:1D with {"classification":0}.
[2019-10-11 00:01:23] (BLE-DATA) 2019/10/11 00:01:23 [E8:44:3E:9D:9B:1D]: {"immobile_state":255}
[2019-10-11 00:01:23] (BLE-DATA) 2019/10/11 00:01:23 [E8:44:3E:9D:9B:1D]: {"classification":0}
[2019-10-11 00:01:23] (ACTIVITY-MONITOR) 2019/10/11 00:01:23 Add 0 to activities.
[2019-10-11 00:01:23] (MainHub) 2019/10/11 00:01:23 //DATA queried by E8:44:3E:9D:9B:1D with {"immobile_state":255}.
[2019-10-11 00:01:23] (MainHub) 2019/10/11 00:01:23 //DATA queried by E8:44:3E:9D:9B:1D with {"classification":0}.
[2019-10-11 00:01:24] (BLE-DATA) 2019/10/11 00:01:24 [E8:44:3E:9D:9B:1D]: {"immobile_state":255}
[2019-10-11 00:01:24] (BLE-DATA) 2019/10/11 00:01:24 [E8:44:3E:9D:9B:1D]: {"classification":0}
[2019-10-11 00:01:24] (ACTIVITY-MONITOR) 2019/10/11 00:01:24 Add 0 to activities.
[2019-10-11 00:01:24] (MainHub) 2019/10/11 00:01:24 //DATA queried by E8:44:3E:9D:9B:1D with {"immobile_state":255}.
[2019-10-11 00:01:24] (MainHub) 2019/10/11 00:01:24 //DATA queried by E8:44:3E:9D:9B:1D with {"classification":0}.
```

Figure 4.1 : Extrait d'un fichier de données brutes

Il est possible de constater une certaine organisation dans ce fichier. En effet, le début de chaque enregistrement contient la date et l'heure d'écriture suivies de du type. Par contre, mise à part la distinction évidente entre la date d'enregistrement et le type d'enregistrement, il n'y a pas de séparateur logique entre les différentes sections d'une même ligne. Mentionnons également qu'un total de 664 fichiers de données brutes étaient disponibles à la fin de la cueillette et devaient être traités.

Comme on peut le remarquer à l'aide de la Figure 4.1, chaque ligne est associée à un type d'enregistrement. Voyons en quoi consiste chacun de ces types :

— ACTIVITY-MONITOR : ce type d'enregistrement est généré par la sous-routine de comptage journalier des activités. Ces données étaient utiles seulement pour effectuer des vérifications lors de la cueillette et ont été éliminées lors du traitement de conversion.

— BLE-DATA : les entrées appartenant à ce type d'enregistrement ont été générées par la sous-routine de communication Bluetooth de l'unité centrale de traitement. Lorsqu'une donnée était transmise ou reçue à l'aide du protocole BLE, une entrée était générée indiquant le moment exact et le contenu du message.

— DAY-ROUTINE : ce type d'enregistrement nous renseigne sur le processus journalier effectué par l'unité centrale de traitement. On y dénote principalement la routine de sauvegarde des données effectuée chaque soir. Bien qu'il soit intéressant pour la validation du bon fonctionnement de la cueillette, ce type est peu utile pour les traitements de détection de changement comportemental.

— HB-ROUTINE : les données sous ce type d'enregistrement proviennent du sous-traitement d'évaluation du statut de connexion. En effet, le statut de connexion entre l'unité centrale de traitement et le serveur de sauvegarde distant était évalué à intervalles régulier. Ces informations ne sont pas retenues pour la suite.

— MainHub : Ce type d'enregistrement provient du processus principal de l'unité centrale de traitement et il regroupe les données de pratiquement tous les processus. De ce fait, il

s'avère que beaucoup de données sont écrites en double dans les fichiers.

— SensorHub : Les données sous ce type d'enregistrement proviennent de la sous-routine de communication entre l'unité centrale de traitement et les capteurs qui communiquent via le réseau WiFi. Les données des capteurs de vitesse ainsi que du dispositif PIR+ se retrouvent sous ce type d'enregistrement.

— Training : lorsque le bracelet est mis en mode exercice, la sous-routine de gestion des entraînements est activée et les entrées générées par celle-ci se retrouvent sous ce type d'enregistrement.

— Sound : un haut-parleur est connecté à l'unité centrale de traitement et tous les sons qui y sont joués génèrent une entrée sous ce type d'enregistrement indiquant quel son a été joué et à quel moment.

4.1.2 REPRÉSENTATION FINALE

Le format brut contient un nombre important d'informations qui ne sont pas toutes pertinentes pour le sujet traité dans ce mémoire et qui peuvent se retrouver en double. Le traitement décrit dans la sous-section qui suit applique un filtre sur les données afin d'en réduire la taille. De plus, les enregistrements ont été formatées pour faciliter la suite du traitement. Au terme du processus de conversion, la représentation finale des données est sous format CSV (colonnes séparées par des virgules) et un seul fichier par personne contenant toutes les journées est disponible. La Figure 4.2 montre un extrait d'un fichier de données traitées.

	A	B	C
1759	2019-10-02 17:52:10	immobile	False
1760	2019-10-02 17:52:10	class	Inactif
1761	2019-10-02 17:52:11	battery	79
1762	2019-10-02 17:52:11	immobile	False
1763	2019-10-02 17:52:11	class	Inactif
1764	2019-10-02 17:52:12	immobile	False
1765	2019-10-02 17:52:12	class	Inactif

Figure 4.2 : Extrait d'un fichier de données traitées

Comme il est possible de la remarquer à l'aide de la Figure 4.2, les fichiers formatés comportent trois colonnes distinctes soit :

1. Le moment (précis à la seconde près) où la donnée a été écrite.
2. La catégorie de donnée.
3. La valeur spécifique de la donnée.

Le Tableau 4.1 liste les catégories de données et les valeurs que peut prendre la dernière colonne du fichier pour chacune de ces catégories.

Tableau 4.1 : Informations contenues dans le fichier de log formaté

Catégorie	Valeures possibles	Description
event	BLEConnectionEvent BLEDisconnectionEvent TrainingEndEvent	Indique qu'un événement spécifique a eu lieu.
offline_classifications	Un Tableau de 5 valeurs numériques	Indique le nombre de chaque activité qui a été reconnue lorsque le bracelet n'était pas connecté avec l'unité centrale de traitement.
immobile	True ou False	Indique si oui ou non le bracelet était immobile au moment inscrit dans le champ date.
class	Inactif Assis-Debout Debout-Assis Marcher Courir	La classe d'activité reconnue par le bracelet au moment inscrit dans le champ date.
charging	True ou False	Indique si oui ou non le bracelet était en cours de chargement au moment inscrit dans le champ date.
battery	Entre 0 et 100	Le pourcentage de charge de la batterie.
speedsensor_kmh	Vitesse en km/h	Vitesse captée par le capteur de vitesse.
sound	Nom du fichier audio joué	Nom du fichier audio joué par l'unité centrale de traitement.
exercise_recognition	Nom de l'exercice reconnu	Exercice reconnu par l'unité centrale de traitement lors d'un entraînement.

On remarque que les données formatées sont bien différentes des données brutes. La section suivante va s'occuper d'expliquer plus en détail le traitement de conversion.

4.1.3 PROCESSUS DE CONVERSION

Le processus de formatage des données brutes s'effectue à l'aide d'un programme conçu spécifiquement pour cette tâche. Le langage de programmation utilisé est le Python, un langage assez populaire pour les traitements de ce genre.

Le programme contient une liste de numéros de participant et itère de manière à traiter les données de chaque participant une à une. Les données sont donc traitées séquentiellement, de la moins récente à la plus récente. Par conséquent, chaque fichier brut est ouvert et traité en ordre chronologique selon sa date de création. Chaque type d'enregistrement est assujetti à un traitement qui lui est propre. Voici les étapes du traitement de conversion :

1. Premièrement, chaque ligne chargée depuis le fichier brut est séparée en trois sections distinctes :
   (a) La date précise à la seconde près de l'enregistrement.
   (b) Le type de l'enregistrement brut
   (c) Les données spécifiques de la catégorie en format brut.
2. Un filtre est ensuite appliqué afin de conserver seulement les types d'enregistrement « MainHub », « Training » et « Sound ». Les enregistrements n'appartenant à aucun de ces types sont donc éliminés.
3. Par la suite, les données spécifiques de chacun des types restants sont redirigées aux bonnes instructions de conversion et y sont formatées. Dans la plupart des cas, la détection de la catégorie de données est effectuée à l'aide de la recherche d'une chaîne

de caractère spécifique. Par exemple, la détection de la chaine de caractère "//RAW-DATA" nous indique que l'enregistrement contient des données d'IMU, le traitement de conversion des valeurs d'IMU en format standard va donc être effectué.

4. Instructions de conversion pour le type d'enregistrement « MainHub » :
   (a) Les données spécifiques appartenant au type d'enregistrement « MainHub » qui sont reconnues à l'aide du mot clé « //RAW-DATA » sont reçues du bracelet lors d'une séquence d'exercices. Les informations d'IMU (6 nombres réels) sont formatées et écrites dans un fichier spécifiquement créé pour stocker ces données.
   (b) Les données spécifiques qui contiennent la chaîne de caractères « BLEConnection » ou « BLEDisconnection » sont générées lors d'une connexion/déconnexion entre le bracelet et l'unité centrale de traitement. À la suite du formatage de cette donnée, le fichier de sortie va contenir une donnée de la catégorie « event » dont la valeur est soit « BLEConnectionEvent » ou « BLEDisconnectionEvent » selon le cas.
   (c) Lorsque la chaîne de caractères « classification_of_the_day » se retrouve dans les données spécifiques du type d'enregistrement « MainHub », nous avons affaire à une donnée générée lorsque le bracelet se reconnecte à l'unité centrale de traitement et retourne les activités effectuées durant la déconnexion. On y retrouve une suite de cinq nombres entiers qui correspondent au comptage de chacune des cinq activités possibles. Ceux-ci sont donc formatés puis écrits dans le fichier standard sous la catégorie « offline_classifications ».
   (d) Les données spécifiques qui sont reconnues à l'aide du mot clé « immobile_state »

nous renseignent si oui ou non le bracelet était en mouvement au moment indiqué par la date et l'heure à la seconde près de l'enregistrement. Les données sont inscrites sous la catégorie « immobile » du fichier de sortie avec la valeur « True » ou « False » selon le cas.

(e) Les données spécifiques qui contiennent la chaîne de caractères « classification » nous renseignent sur l'activité courante reconnue par le bracelet au moment de la génération de l'enregistrement. En format brut, la donnée est un entier qui correspond à l'identifiant de l'activité correspondante. Ce nombre est converti en identifiant textuel (Inactif, Assis-Debout, Debout-Assis, Courir ou Marcher) puis est écrit sous la catégorie « class » du fichier de sortie.

(f) Lorsque la chaîne de caractères « charging_state » se retrouve dans les données spécifiques, nous avons affaire aux renseignements quant à l'état de charge du bracelet. En format brut, la donnée est un nombre entier qui peut prendre la valeur 0 ou 1. Lors du formatage, la valeur écrite est « True » si le bracelet est en cours de chargement et « False » dans le cas contraire. Le tout est finalement écrit sous la catégorie « charging » du fichier de sortie.

(g) Le pourcentage de charge de la batterie se retrouve avec la chaine de caractères « battery » et est écrit sous la catégorie « battery » du fichier de sortie.

(h) Les données spécifiques qui contiennent la chaîne de caractères « speed_kmh » proviennent d'un capteur de vitesse et correspondent à la vitesse enregistrée. Celle-ci est écrite sous la catégorie « speedsensor_kmh ».

5. Instructions de conversion pour le type d'enregistrement « Training ». Les données sous cette catégorie sont générées lorsque le mode entraînement est actif.
   (a) Les données contenant la chaîne « recognition » sont des données générées par l'unité centrale de traitement à la suite de la reconnaissance d'une séquence d'exercices. Les informations de reconnaissance sont enregistrées sous la catégorie « exercise_recognition » du fichier de sortie.
   (b) Une entrée de ce type d'enregistrement qui contient un « training stop » nous renseigne sur le moment de fin de l'entraînement. Cette donnée est enregistrée dans la catégorie « event » du fichier de sortie avec la valeur « TrainingEndEvent ».
6. Instructions de conversion pour la catégorie « sound » : ces données nous renseignent sur le son qui a été joué par l'unité centrale de traitement au moment de l'écriture de l'entrée. On y extrait le fichier audio qui a été joué et le tout est écrit sous la catégorie « sound » du fichier de sortie.

Cela conclut les traitements effectués sur les données brutes. Pour la suite, ces données ne sont plus utilisées et sont remplacées par les données formatées. Ce sont d'ailleurs ces données qui servent d'entrée pour l'extraction d'informations présentée dans la section qui suit.

4.2 EXTRACTION DE L'INFORMATION

Le processus d'extraction de l'information se base sur les données formatées à la suite du processus de conversion afin d'extraire et d'interpréter les informations utiles en vue

d'effectuer la détection de changement comportemental.

Initialement, ce processus était implémenté et exécuté en langage Python. Cependant, le traitement est devenu trop avare en temps de calcul au fil du développement et a été réécrit puis continué en langage C++. La version finale est optimisée afin d'utiliser le moins de ressources processeur et de mémoire possible. De plus, ce traitement s'exécute en parallèle, ce qui réduit encore plus le temps d'exécution.

Comme le processus d'extraction d'informations est relativement complexe, il sera séparé en différentes étapes afin de simplifier sa compréhension. Il est important de noter que ce traitement s'exécute pour un seul participant à la fois ; les explications qui suivent s'appliquent donc au traitement d'un seul participant.

4.2.1 ÉTAPE 1 : CHARGEMENT DES DONNÉES FORMATÉES

Cette étape consiste à charger le fichier formaté du participant traité en mémoire et d'y effectuer un prétraitement.

Premièrement, le fichier est chargé en mémoire dans son entièreté. En effet, les fichiers formatés sont beaucoup plus légers que les fichiers bruts initiaux. En moyenne, un fichier formaté pour un participant contient 180 mégaoctets. Cela permet son chargement en mémoire afin d'économiser du temps lorsque le fichier doit être relu pour la suite.

Chaque ligne du fichier est convertie et chargée dans un conteneur de type « enregistre-

ment ». Voici le format d'un enregistrement :

1. *Timestamp* : l'horodatage, représenté selon le nombre de secondes écoulées depuis le 1er janvier 1970.
2. La clé de donnée (ou catégorie de donnée)
3. La donnée.

Une fois le chargement du fichier effectué, un ensemble contenant tous les enregistrements est stocké en mémoire et facilement accessible. Un prétraitement sur ces données est alors effectué. Ce prétraitement s'occupe de la gestion des déconnexions/reconnexions entre l'unité centrale de traitement et le bracelet et sera décrit dans dans la sous-section qui suit.

4.2.2 ÉTAPE 2 : PRÉTRAITEMENT - GESTION DES DÉCONNEXIONS

Le prétraitement des données formatées a pour but de gérer les déconnexions/reconnexions entre le bracelet et l'unité centrale de traitement. Cette gestion comporte plusieurs enjeux, la présente section va donc s'occuper de présenter plus en détail ces enjeux et les traitements qui sont effectués pour y pallier.

Enjeux

Lorsque le bracelet se déconnecte de l'unité centrale de traitement, les activités reconnues sont enregistrées dans la mémoire du bracelet et sont envoyées en lot lorsque le bracelet se reconnecte sur l'unité centrale de traitement. Comme la mémoire est limitée, seulement le

nombre d'occurrences de chacune des activités est enregistré sans retenir le moment où ces activités ont été effectuées. Cela apporte des problématiques qui doivent être gérées :

1. Normalement, on reçoit, à chaque seconde, un indicateur qui indique s'il y a eu un infime mouvement du bracelet ou non. Cet indicateur est très sensible et il est fréquent que l'activité reconnue soit « Inactif » alors que l'indicateur de mouvement indiquait un mouvement. C'est voulu ainsi, car l'indicateur de mouvement nous renseigne si oui ou non le bracelet est porté. Or, cet indicateur n'est pas disponible lorsqu'il y a une déconnexion, il est donc difficile de savoir si oui ou non la personne portait son bracelet lorsqu'il n'était pas connecté à l'unité centrale de traitement.
2. On ne sait pas quand les activités ont eu lieu ni dans quel ordre. On sait seulement le total cumulé que l'on reçoit lorsque la personne se reconnecte.
3. Normalement, une déconnexion signifie que le bracelet est hors de portée, on associe donc cela à une sortie. Cependant, ce n'est pas toujours le cas. En effet, le bracelet peut devenir hors de portée si l'habitat est trop grand (ex. : s'il y a plusieurs étages dans le domicile), lors d'un problème de connexion, etc.

Ces enjeux complexifient significativement la gestion des déconnexions/reconnexions entre le bracelet et l'unité centrale de traitement. Après analyses et réflexions, des décisions ont été prises afin de pallier à ces enjeux au mieux, ce qui s'est traduit à l'élaboration des traitements ci-bas.

Traitements effectués

Voici les traitements spéciaux qui sont faits actuellement pour gérer les connexions et les déconnexions :

1. Une période de déconnexion qui contient l'activité « Inactif » à 99% et plus est éliminée et remplacée par une période d'inactivité et de non-port du bracelet. Autrement dit, la période de déconnexion est éliminée et transformée en une période régulière démontrant une inactivité complète comme s'il n'y avait pas eu de déconnexion.
2. Une période hors-ligne de moins de 2 minutes n'est vraisemblablement pas une sortie du domicile : un problème de connexion temporaire est plus plausible selon les analyses effectuées. C'est pourquoi elle est éliminée et les activités enregistrées durant la déconnexion sont inscrites comme si elles étaient arrivées normalement en mode en ligne. De plus, si la personne portait son bracelet avant la courte déconnexion, on induit un port de bracelet lors de cette période.
3. Comme mentionné plus tôt, on reçoit le cumulatif des activités effectuées durant la déconnexion une fois que le bracelet se reconnecte. Il incombe de les distribuer dans le passé, mais on ne sait pas quand ni dans quel ordre elles ont été faites. Le traitement qui est fait est donc la séparation uniforme des activités dans le temps. Voici le fonctionnement de cette séparation :
    — On commence par déterminer le nombre total d'activités qui a été effectué en mode hors-ligne. Normalement, le bracelet rapporte l'activité courante toutes

les secondes. On soustrait donc le total des activités effectuées à l'horodatage correspondant au moment où le bracelet s'est reconnecté. Cela nous donne le moment théorique de la déconnexion du bracelet.

— Les analyses ont démontré que le bracelet ne rapporte pas toujours l'activité courante chaque seconde. Afin de trouver la date de début réelle de la déconnexion, la dernière entrée pour laquelle le bracelet était encore connecté à l'unité centrale de traitement est recherchée. Cette recherche est effectuée à partir du début théorique de la déconnexion trouvé à l'étape précédente et l'horodatage de l'enregistrement trouvé est retenu comme date de début de déconnexion. Il est important de noter que cette recherche comporte une limite de temps. En effet, si l'enregistrement n'est pas trouvé dans les N minutes avant le début théorique de la déconnexion, on ne va pas plus loin et c'est le début théorique qui est retenu.

— Finalement, lorsque l'on sait le nombre d'activités effectuées dans chaque catégorie ainsi que le début et la fin de la déconnexion, on peut distribuer ces activités dans les périodes concernées de manière uniforme. À titre d'exemple, si nous avons un cumul de 300 activités « Marche » pour une déconnexion s'échelonnant sur 15 minutes et que nous traitons des périodes de 5 minutes, nous allons distribuer ces 300 activités dans le temps en insérant $300/(15/5) = 100$ occurrences dans chacune des 3 périodes. En plus de cela, un traitement spécial est effectué afin de calculer les bons ratios pour la période où la déconnexion est survenue et pour la période où le bracelet s'est reconnecté.

Ces différents procédés sont effectués à partir des enregistrements chargés à l'étape précédente. Lorsque nécessaire, ceux-ci sont modifiés selon les règles définies ci-haut et finalement utilisées pour l'étape de la compilation de l'information par période.

4.2.3 ÉTAPE 3 : CHARGEMENT DES DONNÉES D'ENTRAÎNEMENT

Les informations sur les entraînements effectués par les participants au cours de la cueillette de données ont été traitées par l'un de mes pairs impliqués dans le projet Acti-DM1. Les données d'IMU recueillies lors des séquences d'entraînement à domicile ont subi une transformation afin de les rendre applicables comme entrée pour certains algorithmes d'apprentissage machine. À la suite de ces différents procédés, un fichier en format JSON a été généré. La Figure 4.3 montre un extrait d'un fichier contenant les informations sur les entraînements.

Figure 4.3 : Extrait d'un fichier d'information sur les entraînements

Les informations contenues dans ce type de fichier apparaissent sous une arborescence dont le premier niveau liste les numéros de participants. Pour chaque participant, le second niveau de l'arborescence comporte les dates pour lesquelles des informations sont disponibles. Le troisième niveau, quant à lui, comporte la plupart des informations utiles. Voici les informations que l'on retrouve pour ce niveau :

1. Prescriptions : liste des différentes séquences d'exercices qui devaient être effectuées par l'individu dans la journée. L'information est organisée sous forme de liste où l'on retrouve les types d'exercices (ex. : fente, squat, assis-debout) de même que le nombre de séries et leurs répétitions (ex. : 2 répétitions d'une série de 10 squats).
2. Rappelé : indique si oui ou non un rappel sonore a été effectué dans le cas où des exercices étaient prescrits pour la journée.
3. Entraînements : liste les séquences d'entraînement réellement effectuées. Le déroule-

ment de l'entraînement est alors représenté à l'aide d'une séquence d'étapes comprenant pour chacune le type d'exercice effectué ainsi que le type d'exercice attendu. On retrouve aussi l'heure de début et de fin de l'entraînement et son indicateur de validité.

Dans le cadre de la compilation des données par période, on s'intéresse seulement aux moments où un entraînement était en cours. Le chargement des données d'entraînement s'intéresse donc à extraire tous les intervalles de temps pour lesquels un entraînement valide a été effectué. Pour ce faire, tous les entraînements sont lus et l'horodatage du début et de la fin de chacune des séquences valides est gardé en mémoire.

4.2.4 ÉTAPE 4 : COMPILATION DE L'INFORMATION PAR PÉRIODE

Cette étape consiste à regrouper les données par période afin d'extraire l'information utile pour chaque intervalle de temps (ou période). Une période est un espace de temps dont la durée est définie et pour laquelle les données sont comptabilisées et regroupées. Il sera d'abord question de présenter les caractéristiques extraites pour chaque période au terme de ce traitement pour ensuite enchaîner avec le fonctionnement de l'extraction tout en terminant avec les détails des calculs effectués pour chaque caractéristique.

Caractéristiques extraites

Avant d'attaquer le fonctionnement du traitement de regroupement, il incombe de présenter les caractéristiques extraites durant le processus. Le Tableau 4.2 liste les caractéristiques

extraites lors du processus d'extraction.

Tableau 4.2 : Caractéristiques extraites lors du processus d'extraction de l'information

Caractéristique	Description
wear_percent	Pourcentage de l'intervalle de temps pour lequel le bracelet était porté.
active_percent	Pourcentage de l'intervalle de temps pour lequel une activité autre qu'Inactif était en cours.
offline_activity_percent	Pourcentage de l'intervalle de temps pour lequel le bracelet n'était pas connecté à l'unité centrale de traitement.
no_acti_pct	Pourcentage de l'intervalle de temps pour lequel l'activité Inactif était en cours.
low_acti_pct	Pourcentage de l'intervalle de temps pour lequel les activités Assis-Debout et Debout-Assis étaient en cours.
moderate_acti_pct	Pourcentage de l'intervalle de temps pour lequel les activités Marcher et Courir étaient en cours.
no_offl_acti_pct	Pourcentage de l'intervalle de temps en mode hors-ligne pour lequel l'activité Inactif était en cours.
low_offl_acti_pct	Pourcentage de l'intervalle de temps en mode hors-ligne pour lequel les activités Assis-Debout et Debout-Assis étaient en cours.
moderate_offl_acti_pct	Pourcentage de l'intervalle de temps en mode hors-ligne pour lequel les activités Marcher et Courir étaient en cours.
lost_entries_percent	Pourcentage de l'intervalle pour lequel aucune donnée n'est disponible.

Les caractéristiques présentées dans le tableau ci-dessus vont servir à des fins de validation visuelle ainsi que pour l'interface web Acti-DM1. Parallèlement à la génération de ces caractéristiques, une version quelque peu différente est aussi générée et conçue spécifiquement pour les algorithmes de détection de changements comportementaux. Le Tableau 4.3 expose ces caractéristiques spécifiques.

Tableau 4.3 : Caractéristiques spécifiquement extraites pour les algorithmes de changement comportemental

Caractéristique	Description
wear_seconds	Nombre de secondes dans l'intervalle de temps pour lequel le bracelet était porté.
active_seconds	Nombre de secondes dans l'intervalle de temps pour lequel une activité autre qu'Inactif était en cours.
low_activity_seconds	Nombre de secondes dans l'intervalle de temps pour lequel les activités Assis-Debout et Debout-Assis était en cours.
moderate_activity_seconds	Nombre de secondes dans l'intervalle de temps pour lequel les activités Marcher et Courir étaient en cours.
outside_seconds	Nombre de secondes dans l'intervalle de temps pour lequel le bracelet n'était pas connecté à l'unité centrale de traitement.
outside_active_seconds	Nombre de secondes dans l'intervalle de temps pour lequel le bracelet n'était pas connecté à l'unité centrale de traitement, mais que des activités autres qu'Inactif ont été enregistrées.
lost_seconds	Nombre de secondes dans l'intervalle de temps pour lequel aucune donnée n'est disponible.

Afin de mieux situer les différents processus de traitement d'informations et les données qui y sont traitées, la Figure 4.4 effectue une synthèse des traitements effectués permettant de transformer les données brutes vers les caractéristiques présentées ci-haut.

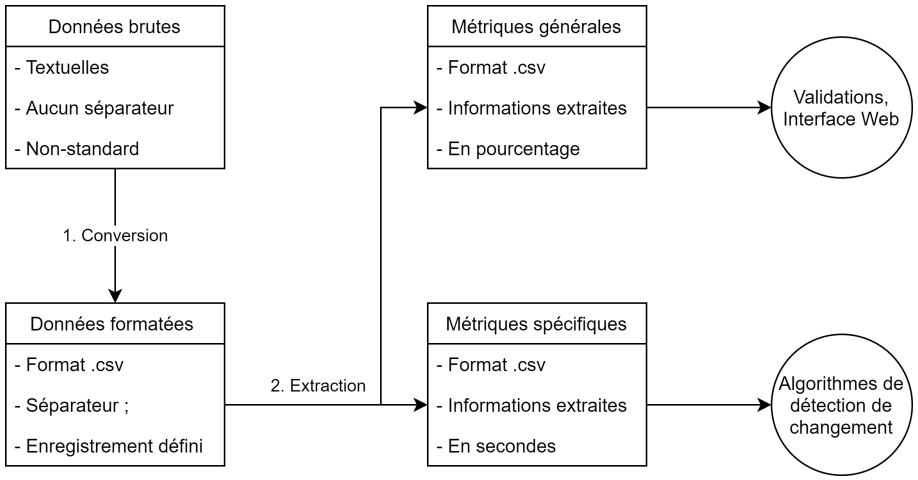


Figure 4.4 : Synthèse des traitements effectués sur les données

Pour poursuivre, la sous-section suivante va présenter les différentes manipulations qui ont permis d'extraire l'information présentée dans les Tableaux 4.2 et 4.3 à partir des données prétraitées stockées en mémoire.

Processus d'extraction

Le traitement de comptabilisation de l'information par période prend en paramètre la durée de l'intervalle de temps en minutes et la date de début de l'expérimentation. Pour les besoins de ce projet, deux traitements par participant ont été effectués. Le premier pour des périodes de 5 minutes et le deuxième pour des périodes d'une journée.

En commençant par la date de début spécifique au participant courant, les caractéristiques sont calculées une période à la fois de façon séquentielle et le traitement se termine lorsque la fin des données est atteinte.

La première étape du processus d'extraction est de consolider les informations d'activités et de mouvements pour la période traitée. Pour ce faire, tous les enregistrements préalablement chargés qui sont associés à la période en cours de traitement sont lus un à la fois et les opérations suivantes y sont effectuées :

1. À chaque fois qu'une entrée de l'indicateur de mouvement est détectée, un compteur est incrémenté afin de garder en mémoire le nombre de fois que le bracelet a rapporté le statut de mouvement durant la période. Normalement, si le bracelet était connecté à l'unité centrale de traitement pour toute la durée de l'intervalle, ce compteur devrait correspondre à la durée de l'intervalle de temps en secondes une fois tous les enregistrements lus.
2. Le nombre de fois que l'indicateur de mouvement indique qu'il y a mouvement est sauvegardé dans un compteur.
3. À chaque fois qu'une entrée de l'indicateur d'activité est détectée, que ce soit en ligne ou hors-ligne, un compteur est incrémenté.
4. Les occurrences de chacune des activités possibles (Inactif, Assis-Debout, Debout-Assis, Courir ou Marcher) sont comptées, y compris les activités qui ont été effectuées en mode hors-ligne. L'addition de ces cinq compteurs va être égale aux compteurs du nombre d'activités rapporté décrit ci-haut.
5. Cinq compteurs comprenant spécifiquement les occurrences de chacune des activités possibles effectuées en mode hors-ligne sont aussi incrémentés.

La deuxième étape du processus d'extraction consiste à retrouver les entraînements effectués durant la période courante, s'il y a lieu, et de compter le nombre de secondes pour lesquelles cet entraînement était en court.

La dernière étape consiste à calculer les caractéristiques à l'aide des informations recueillies jusqu'à maintenant.

Calculs effectués

Avant d'expliquer les différents calculs effectués, voici un résumé de l'information disponible avant les calculs :

— Horodatage du début et de la fin de l'intervalle traité ainsi que sa durée (la durée est définie en tant que *periodDuration*).

— Compteur du nombre de statuts de mouvement rapportés durant la période, défini en tant que *liveStateEntries* pour la suite.

— Compteur du nombre de statuts de mouvement rapportés durant la période dont l'indicateur rapporte qu'un mouvement est en cours, défini en tant que *liveMovingStateEntries* pour la suite.

— Compteur du nombre total d'activités rapportées en mode en ligne durant la période, défini en tant que *liveActivityEntries* pour la suite.

— Compteur du nombre total d'activités rapportées en mode hors ligne durant la période, défini en tant que $offlineActivityEntries$ pour la suite.

— Nombre de secondes pour lesquelles un entraînement était en cours durant la période, défini en tant que $trainingSeconds$ pour la suite.

— Comptage du nombre d'activités rapportées de chaque type. Le résultat est un regroupement de cinq compteurs (soit un compteur par type d'activité) qui nous renseignent combien de fois chaque activité a été rapportée durant la période, définis en tant que $classesEntries(X)$ où X est l'identifiant d'activité.

— Les compteurs $classesEntries(X)$ définis ci-haut comptent toutes les activités, qu'elles aient été effectuées en mode en ligne ou en mode hors ligne. Une autre donnée disponible compte les activités effectuées en mode hors-ligne seulement et est définie en tant que $offlineClassesEntries(X)$ où X est l'identifiant d'activité.

Maintenant que les informations de base sont définies, il est temps de présenter les opérations effectuées pour chaque caractéristique.

Tout d'abord, des variables temporaires sont calculées afin de simplifier les opérations mathématiques pour le calcul des caractéristiques.

La première variable temporaire indique le nombre total d'activités rapporté durant la période courante. Le nombre de secondes d'entraînement y est inclus, car, même si les activités ne sont pas rapportées durant un entraînement, la personne est active et le bracelet est connecté à l'unité centrale de traitement.

$$totalActivityEntries = liveActivityEntries + offlineActivityEntries + trainingSeconds$$

La deuxième variable compte le nombre d'entrées pour lesquelles la personne qui portait le bracelet était active durant la période.

$$activeEntries = classesEntries(assisDebout) + classesEntries(deboutAssis) + classesEntries(marcher) + classesEntries(courir) + trainingSeconds$$

La troisième variable estime le nombre d'entrées perdues pour la période traitée. Le bracelet est censé rapporter son activité toutes les secondes sauf en mode d'entraînement. Lorsque ce n'est pas le cas, les secondes pour lesquelles aucune activité n'est rapportée sont considérées comme perdues.

$$lost_entries = periodDuration - liveActivityEntries - offlineActivityEntries - trainingSeconds$$

La quatrième variable temporaire indique le pourcentage des entrées d'indicateur de mouvement qui indiquent qu'un mouvement a été détecté. La formule prend en considération que les entraînements induisent qu'il y a mouvement et que l'indicateur de mouvement n'est pas rapporté lors de ceux-ci.

$$movingPercent = \frac{liveMovingStateEntries + trainingSeconds}{liveStateEntries + trainingSeconds}$$

Les dernières variables temporaires convertissent le nombre d'entrées de chaque activité en pourcentage.

$$classesPercent(X) = \frac{classesEntries(X)}{totalActivityEntries}$$

$$offlineClassesPercent(X) = \frac{offlineClassesEntries(X)}{offlineActivityEntries}$$

Voici maintenant les calculs effectués pour créer les caractéristiques générales apparaissant dans le Tableau 4.2.

Le calcul pour la première caractéristique convertit simplement le nombre d'entrées

perdues en pourcentage.

$$lost_entries_percent = \frac{lost_entries}{periodDuration}$$

Le pourcentage de port du bracelet se base sur le pourcentage de mouvement calculé précédemment. Cependant, un facteur d'ajustement a lieu afin de pénaliser le résultat dans le cas où beaucoup de données perdues sont détectées dans la période courante. En effet, lorsque le bracelet ne rapporte pas son activité, le bracelet n'est pas considéré comme porté.

$$wear_percent = movingPercent\left(1 - \frac{lost_entries}{periodDuration}\right)$$

Le calcul pour cette caractéristique convertit en pourcentage le nombre d'entrées indiquant qu'une activité autre qu'Inactif était en cours.

$$active_percent = \frac{activeEntries}{totalActivityEntries}$$

Enfin, le nombre d'entrées générées en mode hors ligne est converti en pourcentage.

$$offline_activity_percent = \frac{offlineActivityEntries}{totalActivityEntries}$$

Ces caractéristiques regroupent les activités reconnues en sous-catégories. Les activités

Assis-Debout et Debout-Assis sont regroupées sous la catégorie des activités faibles alors que les activités Marcher et Courir sous la catégorie des activités modérées.

$$
\begin{aligned}
no_acti_pct &= classesPercent(inactif) \\
low_acti_pct &= classesPercent(assisDebout) \\
&\quad + classesPercent(deboutAssis) \\
moderate_acti_pct &= classesPercent(marcher) + classesPercent(courir) \\
no_offl_acti_pct &= offlineClassesPercent(inactif) \\
low_offl_acti_pct &= offlineClassesPercent(assisDebout) \\
&\quad + offlineClassesPercent(deboutAssis) \\
moderate_offl_acti_pct &= offlineClassesPercent(marcher) \\
&\quad + offlineClassesPercent(courir)
\end{aligned}
$$

Les calculs effectués pour obtenir les caractéristiques exposées dans le Tableau 4.3 comportent principalement une conversion du pourcentage en nombre de secondes de certaines caractéristiques générales définies précédemment.

$$
\begin{aligned}
wear_seconds &= periodDuration \times wear_percent \\
active_seconds &= periodDuration \times active_percent
\end{aligned}
$$

$$low_activity_seconds = periodDuration \times low_acti_pct$$

$$moderate_activity_seconds = periodDuration \times moderate_acti_pct$$

$$outside_seconds = periodDuration \times offline_activity_percent$$

$$lost_seconds = periodDuration \times lost_entries_percent$$

$$outside_active_seconds = periodDuration \times (low_offl_acti_pct + moderate_offl_acti_pct)$$

4.3 SOMMAIRE

En somme, ce chapitre a présenté les différentes opérations qui ont été effectuées sur les données afin de les préparer pour l'analyse des changements comportementaux. Les fichiers originaux non standards générés par l'unité centrale de traitement ont été traités par un premier programme informatique qui a permis de filtrer et de classifier les données, ce qui a donné de nouveaux fichiers de type standards. Ces fichiers ont servi d'entrée au deuxième programme qui avait pour objectif l'extraction de l'information utile. Ce programme a permis de compiler l'information par période à l'aide de différentes manipulations et calculs, ce qui a finalement mené à la création des fichiers qui serviront pour l'étude des changements comportementaux.

CHAPITRE V

DÉTECTION DE CHANGEMENT COMPORTEMENTAL

Le chapitre précédent a démontré les différents traitements effectués sur les données brutes afin de produire un fichier par participant qui va servir d'entrée pour le processus de détection de changement comportemental. Le présent chapitre s'appuie sur ces informations afin de poursuivre avec la détection de changement comportemental effectuée dans le cadre de ce mémoire. Pour commencer, les différentes étapes qui ont été appliquées de même que le programme informatique développé à cet égard seront expliqués. Ensuite, les résultats obtenus à la suite du processus de détection de changement comportemental seront exposés. Finalement, une analyse approfondie des résultats obtenus sera effectuée, suivie d'une discussion.

5.1 MÉTHODE

L'étape de la détection de changements comportementaux se base sur les fichiers qui contiennent les caractéristiques spécifiquement extraites pour les algorithmes de détection de changements comportementaux dont le contenu a été rapporté dans le Tableau 4.3. D'ailleurs, un programme informatique, développé en langage Python, prend en entrée ces fichiers et applique les différents algorithmes de détection de changement comportemental choisis (sw-PCAR et RulSIF). La présente section va s'occuper de décrire la méthode utilisée pour effectuer la détection de changement comportemental et la structure du programme informatique développé.

5.1.1 MODES DE COMPARAISON

Les algorithmes de détection de changement comportemental sont appliqués sur deux fenêtres de données de même taille qui doivent être comparées afin d'en ressortir un score de changement. Un outil de sélection des fenêtres est intégré au programme informatique et permet une certaine latitude quant à la règle de sélection des fenêtres pour chaque cycle de comparaison.

Dans un premier temps, il incombe de spécifier la taille (en nombre de jours) d'une fenêtre. Pour les besoins du présent projet, la taille des fenêtres est d'une semaine (7 jours). Ensuite, il est possible de décider le mode de comparaison des fenêtres. Soit la première fenêtre est fixe (ex. : elle correspond à la première semaine de l'expérimentation), soit elle glisse dans le temps. En fait, il est possible de spécifier de combien de jours on glisse chacune des fenêtres dans le temps à chaque cycle de comparaison. Pour que la première fenêtre soit fixe, il suffit d'indiquer zéro glissement pour la première fenêtre et c'est d'ailleurs ce qui a été fait dans le cadre du projet Acti-DM1. La deuxième fenêtre, quant à elle, glisse de 7 jours à chaque cycle. Avec cette configuration, la première semaine est comparée avec toutes les autres. Par exemple, si nous avons 10 semaines de données pour un participant, 9 comparaisons seront effectuées : la semaine 1 avec la semaine 2, la semaine 1 avec la semaine 3, la semaine 1 avec la semaine 4 et ainsi de suite.

La Figure 5.1 résume la méthode de comparaison des fenêtres. Celle-ci représente le processus de sélection des fenêtres pour un nombre de jours égal à 70 (10 semaines). La

semaine 1, qui est la première fenêtre, est statique et va toujours comporter les données des 7 premiers jours. La deuxième fenêtre est aussi d'une taille de 7 jours (1 semaine) et va glisser de 7 jours à chaque cycle de comparaison. De ce fait, la première comparaison s'effectue entre la semaine 1 et la semaine 2. Ensuite, la fenêtre 2 glisse de 7 jours et contient alors les données de la semaine 3. La deuxième comparaison effectuée sera donc faite entre les données de la première semaine et les données de la troisième semaine. Ce processus se répète jusqu'à la fin des données disponibles pour le participant traité.

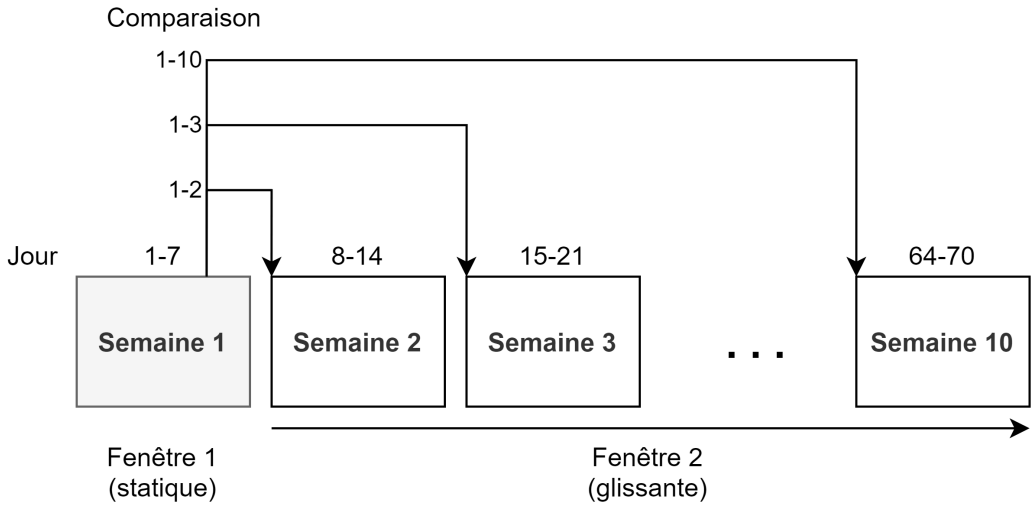


Figure 5.1 : Synthèse de la méthode de sélection des fenêtres pour comparaison

5.1.2 CHARGEMENT DES CARACTÉRISTIQUES

Avant de pouvoir calculer un score de changement entre deux fenêtres de données, il est tout d'abord nécessaire d'extraire et de représenter les données de chacune des fenêtres correctement. C'est pourquoi une fonction de chargement des données d'une fenêtre est

intégrée au programme informatique de détection de changement comportemental. Cette fonction charge les enregistrements du fichier contenant les caractéristiques entre le jour de début et le jour de fin spécifié. En outre, il retourne un ensemble d'enregistrements contenant les données des caractéristiques pour la période chargée.

Les caractéristiques sélectionnées lors du chargement des caractéristiques pour une semaine ont été précédemment décrites dans le Tableau 4.3. Voici ceux qui sont chargés :

— wear_seconds

— low_activity_seconds

— moderate_activity_seconds

— outside_seconds

— outside_active_seconds

— lost_seconds

De ce fait, la fonction de chargement est appelée une fois pour la première semaine (jours 1 à 7) et les caractéristiques de celle-ci sont sauvegardées. Ensuite, à chaque cycle de comparaison, la deuxième fenêtre est chargée et les informations recueillies du fichier pour la première et deuxième fenêtre sont utilisées pour l'exécution des algorithmes de détection de changements.

5.1.3 ALGORITHMES

Comme il a été mentionné plus tôt, pour chaque cycle de comparaison, les caractéristiques sont disponibles pour chacune des deux fenêtres et sont envoyées aux algorithmes de détection de changement comportemental. Sauf pour une méthode spécifique de RulSIF, les algorithmes comparent seulement une caractéristique à la fois. Cela implique que chacun des algorithmes sera exécuté autant de fois qu'il y a de caractéristiques. Par exemple, les informations issues de la caractéristique wear_seconds seront comparées entre la fenêtre 1 et la fenêtre 2, ce qui va donner un premier score de changement relatif à cette caractéristique. Ensuite, ce sera la caractéristique low_activity_seconds qui fera l'objet d'une comparaison entre les deux fenêtres, ce qui donnera un nouveau score de changement. Ce processus s'effectue pour toutes les caractéristiques disponibles. Aussi, certaines manipulations supplémentaires sont effectuées sur les informations envoyées aux algorithmes de détection de changement comportemental afin de préparer l'entrée selon les besoins spécifiques de chacun. Voyons les détails spécifiques de chacun des algorithmes utilisés.

sw-PCAR

Au moment de l'élaboration du programme informatique utilisé pour la détection de changement, l'implémentation de l'algorithme sw-PCAR n'était pas disponible dans la communauté du logiciel libre. L'écriture de cet algorithme a donc été effectuée à partir des informations contenues dans l'article de Sprint *et al.* (2016b) afin d'être intégré au programme informatique. Cette implémentation encapsule le fonctionnement de l'algorithme sw-PCAR

tel que décrit dans le Chapitre 2. En entrée, il suffit de spécifier le contenu des deux fenêtres à comparer et l'algorithme va retourner le score de changement et un booléen qui sera à vrai si un changement conséquent a été détecté. Rappelons que sw-PCAR consiste en une modification de PCAR qui peut entre autres accepter des caractéristiques sous la forme de nombres réels. C'est pourquoi il est possible de lui envoyer directement le contenu des fenêtres qui sont déjà sous cette forme.

RulSIF

Contrairement à sw-PCAR, l'implémentation de l'algorithme RulSIF était disponible en version libre [15] au moment de l'élaboration du programme informatique, ce qui a grandement simplifié la mise en place de celui-ci. Cette implémentation permet une certaine latitude quant aux paramètres d'exécution. En effet, il est possible de spécifier certains paramètres afin d'optimiser son exécution selon la situation. Par exemple, pour les besoins du projet, le paramètre $alpha$ (voir section 2.1.3) est configuré à 0.1 lors de l'exécution des comparaisons avec cet algorithme. Cette valeur a été déterminée expérimentalement et a été retenue selon le critère des bonnes pratiques et de la stabilité des résultats.

Les tests préliminaires effectués ont démontré une grande variance des résultats même lorsque l'algorithme est exécuté plusieurs fois avec les mêmes données. Cela est dû à la nature probabiliste de la méthode et la solution développée pour pallier à cela est l'utilisation d'une moyenne. En effet, RulSIF est exécuté N fois avec les deux mêmes fenêtres et la moyenne des

15. https://github.com/hoxo-m/densratio_py

scores de changement obtenue est calculée puis retournée. Pour les besoins de ce projet, le paramètre N est égal à 100.

L'exécution de RulSIF est donc encapsulée dans une routine qui accepte les deux fenêtres à comparer ainsi que le mode d'exécution souhaité en entrée et retourne le score de changement moyen en sortie. Les fenêtres de données sont converties dans le bon format selon le mode d'exécution choisi (vecteur, matrice ou matrice de caractéristiques). Ces modes permettent d'appliquer RulSIF de manière différente sur les mêmes données selon leur arrangement.

Pour le mode vecteur, RulSIF effectue la comparaison de deux vecteurs qui correspondent à une distribution de probabilité calculée à partir de chacune des fenêtres de données. Pour ce faire, les deux fenêtres sont tout d'abord converties en fenêtres agrégées selon le même principe que pour sw-PCAR : l'ensemble des journées comprises dans une fenêtre sont agrégées de façon à transformer cet ensemble en seulement une journée dont la valeur de chaque intervalle correspond à la moyenne de cet intervalle pour l'ensemble des journées avant agrégation. Le tout est converti en distribution de probabilités selon une méthode assez simple :

— On ajoute la valeur 1 à chaque élément du vecteur agrégé afin d'éliminer tous les zéros, ce qui permet d'éviter les divisions par zéro pour la suite.

— La somme de tous les éléments de chaque vecteur agrégé est calculée. Le résultat est noté en tant que $total$.

— On transforme chaque élément du vecteur agrégé en un nombre inférieur à 1. Pour

ce faire, chaque élément est divisé par la somme de tous les éléments noté $total$. La transformation suivante est donc effectuée pour chacun : $element = \frac{element}{total}$

La Figure 5.2 représente un vecteur contenant une distribution de probabilités créée spécifiquement pour la méthode RulSIF-Vecteur. La taille du vecteur est égale au nombre de périodes contenues dans une journée (ex. : un vecteur de taille 24).

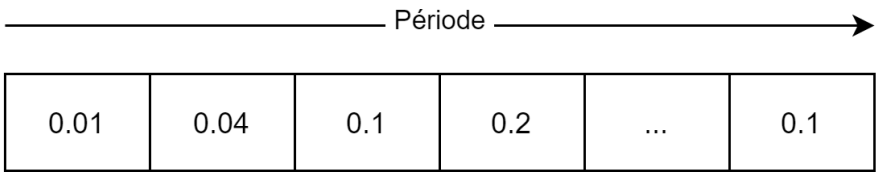


Figure 5.2 : Vecteur de probabilités créé pour la méthode RulSIF-vecteur

Le mode matrice conserve le format des fenêtres d'entrées et convertit chaque journée (ligne de la matrice) en distribution de probabilité selon le même procédé que décrit plus haut avec le mode de fonctionnement de type vecteur.

La Figure 5.3 représente bien cette matrice. On peut voir que chaque ligne de celle-ci correspond à une journée et que le nombre de lignes correspond à la taille de chaque fenêtre.

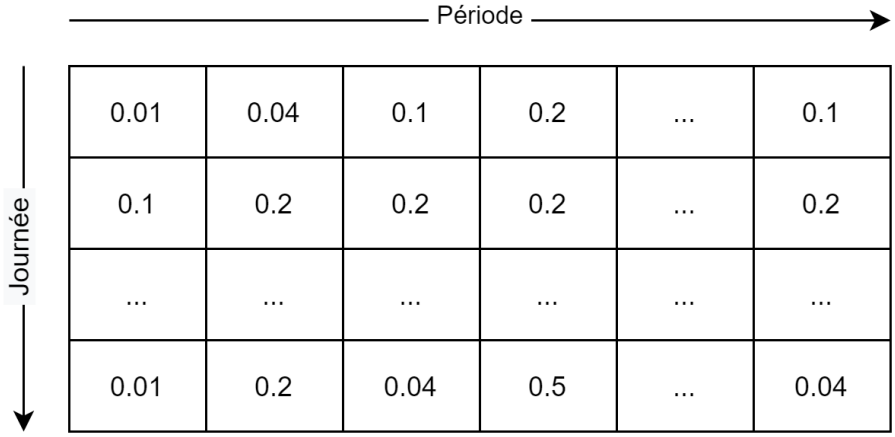


Figure 5.3 : Matrice de probabilités créée pour la méthode RulSIF-matrice

Pour le mode d'exécution matrice de caractéristiques de RulSIF, un vecteur est créé pour chaque caractéristique de chaque fenêtre de la même manière que pour la méthode RulSIF-Vecteur. Contrairement à la méthode RulSIF-Vecteur qui effectue l'exécution de RulSIF une fois par caractéristique, cette méthode combine les vecteurs de toutes les caractéristiques en une matrice et exécute RulSIF une seule fois avec ces données. Ce mode d'exécution va donc générer un seul score de changement par semaine, mais celui-ci va prendre en compte toutes les caractéristiques.

La Figure 5.4 démontre de quelle façon les vecteurs agrégés de probabilités tels que représentés dans la Figure 5.2 ont été empilés pour former une matrice.

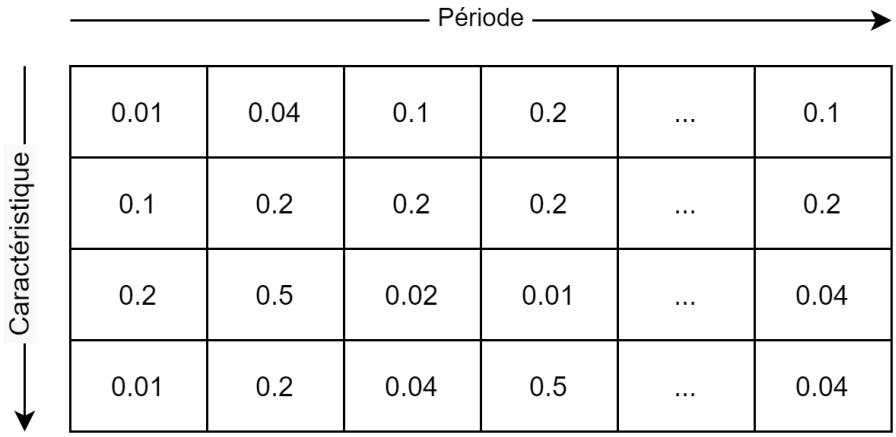


Figure 5.4 : Matrice de probabilités créée pour la méthode RulSIF-caractéristiques

5.1.4 SORTIE

L'algorithme sw-PCAR et les trois variantes de RulSIF totalisent quatre modes de détection de changement pour lesquels des scores de changement sont générés. Pour chacun de ces modes, un fichier en format CSV est généré et contient les scores de changement pour chaque participant. Ensuite, ces fichiers sont utilisés afin de créer des graphiques qui représenteront les informations de manière plus visuelle. Avant d'expliquer le format des fichiers et du graphique, il est important de bien comprendre le principe de numérotation des jours et des semaines.

Identification des journées et des semaines

Les algorithmes utilisent une méthode d'identification des jours et des semaines selon une numérotation. Comme chaque participant a débuté l'expérimentation à un moment dif-

férent, une date de début est associée à chacun. Cette date indique le moment officiel où la personne a commencé à prendre part à la cueillette de manière quotidienne. Celle-ci marque le jour numéro 1 de la semaine numéro 1. De ce fait, la semaine 1 va correspondre aux jours 1 à 7 et ainsi de suite. Lorsque les résultats seront présentés dans les sections qui suivent, la date de début propre à chacun sera spécifiée.

Fichiers de sortie

À chaque fois qu'une méthode de détection de changement est exécutée, un fichier portant le nom de la méthode et dont le contenu est en format CSV est créé afin de stocker les résultats obtenus. Au terme de l'exécution des quatre méthodes, les fichiers suivants sont créés :

— swPCAR_CS.csv : contient les scores de changement de la méthode sw-PCAR

— RulSIF-vect_CS.csv : contient les scores de changement pour le mode de fonctionnement vecteur de la méthode RulSIF.

— RulSIF-mat_CS.csv : contient les scores de changement pour le mode de fonctionnement Matrice de la méthode RulSIF.

— RulSIF-mat-fea_CS.csv : contient les scores de changement pour le mode de fonctionnement Matrice de caractéristiques de la méthode RulSIF.

La Figure 5.5 montre un exemple d'un fichier de sortie. L'algorithme représenté est sw-PCAR, mais la structure est identique pour les trois modes de RulSIF. Pour chaque participant, il est possible de voir le score de changement de chaque comparaison effectuée, et ce pour

toutes les caractéristiques disponibles.

Participant	Feature	1-2	1-3	1-4	1-5	1-6	1-7	1-8	1-9
481	wear_seconds	0.011	0.217	3.201	3.201	0.9263	0.6411	0.5724	0.7447
481	low_activity_seconds	0.1786	0.1166	2.7411	2.7411	1.6084	1.546	1.4009	1.7422
481	moderate_activity_seconds	0.1311	0.1335	2.1824	2.1824	1.7965	1.4854	1.3601	1.4967
481	outside_seconds	0.0	0.0	0.0	0.0	4.7994	0.0	0.0	0.0
481	outside_active_seconds	0.0	0.0	0.0	0.0	1.4324	0.0	0.0	0.0
907	wear_seconds	0.3504	0.8433	0.9053	0.6717	0.9441	1.3879	0.999	0.8537
907	low_activity_seconds	0.125	0.3138	0.3185	0.2297	0.4834	0.9154	0.5889	0.5823

Figure 5.5 : Exemple d'un fichier de sortie généré à la suite de la génération des scores de changement

La première colonne du fichier indique le numéro de participant, la deuxième indique la caractéristique pour laquelle les comparaisons ont été effectuées et les autres colonnes nous renseignent sur le score de changement obtenu pour chaque cycle de comparaison effectuée. Par exemple, la colonne identifiée avec le libellé 1-2 renferme le score de changement calculé entre la semaine 1 et la semaine 2, le libellé 1-3 indique le résultat de comparaison entre la semaine 1 et 3 et ainsi de suite.

Ces fichiers sont ensuite utilisés afin de générer des graphiques qui permettront de représenter les résultats de manière plus visuelle.

Graphiques

Les graphiques seront présentés à la section suivante, mais il sera tout d'abord nécessaire d'expliquer leur fonctionnement pour faciliter leur compréhension.

Le premier type de graphique expose le nombre de secondes passé chaque jour à effectuer

les actions suivantes :

— Porter le bracelet.

— Effectuer une activité dont l'effort est faible.

— Effectuer une activité dont l'effort est modéré.

L'axe des abscisses correspond au numéro de la journée et l'axe des ordonnées correspond au nombre de secondes. La Figure 5.6 est un graphique de ce type.

Le deuxième type de graphique présente les scores de changement obtenus à l'aide d'une méthode de détection de changement comportemental spécifique. Pour ce type de graphique, l'axe des abscisses correspond au cycle de comparaison et l'axe des ordonnées correspond au score de changement obtenu. La Figure 5.7 qui sera présentée dans la section suivante est un graphique de ce type.

5.2 RÉSULTATS DE L'ANALYSE DES CHANGEMENTS COMPORTEMENTAUX

Les résultats obtenus suivant la cueillette de données et l'étude des changements comportementaux seront tout d'abord présentés individuellement pour chaque participant ayant participé à la cueillette de façon régulière. Sur les 10 participants qui avaient l'aide de la technologie, seulement 5 personnes ont participé à la cueillette jusqu'à la fin. Sur les cinq personnes qui n'ont pas complété la cueillette, une personne a été exclue à cause de difficultés avec sa connexion internet, deux personnes se sont retirées à la suite d'un décès d'un membre de leur famille, une personne s'est retirée pour des raisons familiales et une personne s'est

retirée à la suite d'une chirurgie.

5.2.1 PARTICIPANT 481

Le participant numéro 481 a commencé la cueillette en date du 3 octobre 2019 et les données de neuf semaines complètes ont pu être traitées. Aucune information notable n'était inscrite dans le carnet de bord. En effet, aucun événement de vie n'a été rapporté et peu de sorties du domicile ont été indiquées. Un bris matériel est survenu près de la quatrième semaine et la technologie a pu être remise opérationnelle seulement vers la fin de la cinquième semaine.

La Figure 5.6 contient un graphique qui illustre le temps écoulé pour chaque catégorie d'action effectuée. On y perçoit facilement la panne matérielle qui est survenue entre la semaine 4 et la semaine 5. En effet, la courbe indiquant la non-disponibilité des données du bracelet est totale pour ces deux semaines.

On perçoit aussi un port du bracelet moins constant pour les semaines 6, 8 et 9. Il est possible de remarquer que le bracelet n'est pas porté les journées où la courbe du port du bracelet est nulle et qu'aucune donnée manquante n'est rapportée.

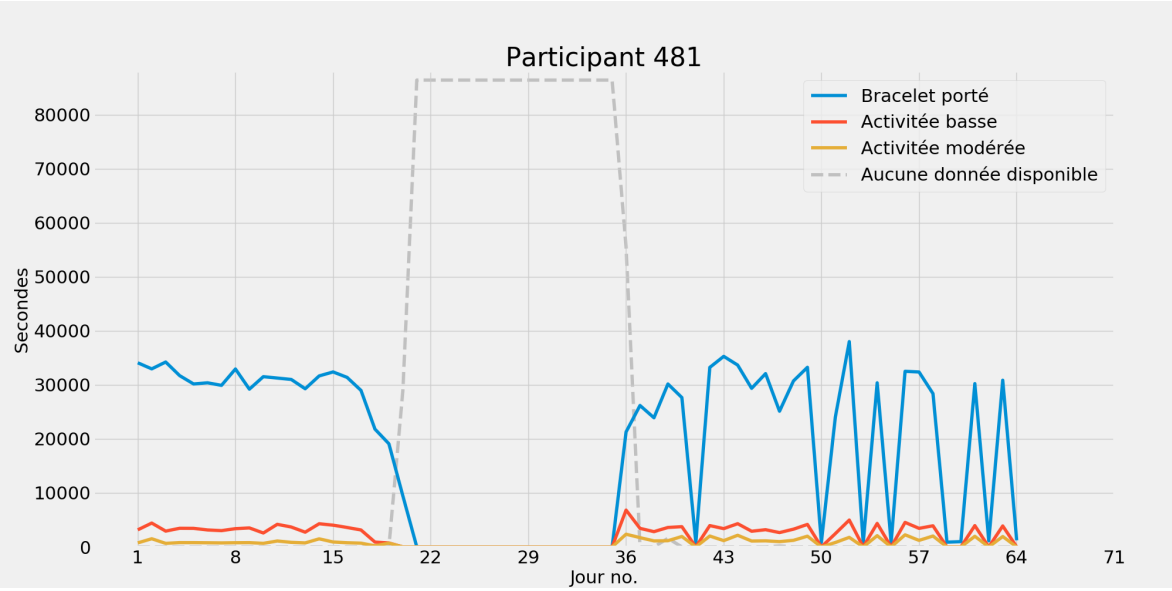


Figure 5.6 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 481

Détection de changement comportemental

Les algorithmes sw-PCAR et les trois modes d'exécution de l'algorithme RulSIF ont été appliqués aux données des neuf semaines, ce qui permet de calculer huit scores de changement pour chaque algorithme. La Figure 5.7 expose les scores de changement obtenus à l'aide de l'algorithme sw-PCAR et les Figures 5.8, 5.9 et 5.10 exposent les résultats de chacun des modes d'exécution de RulSIF.

Le participant numéro 481 n'a subi aucun événement de vie notable qui aurait pu se refléter dans les données de changements. Bien évidemment, la panne matérielle a été rapportée de façon explicite par tous les algorithmes de détection de changements. L'algorithme le plus

sensible au port du bracelet moins constant après la panne matérielle est RulSIF-matrice. On remarque qu'il détecte efficacement les irrégularités de la semaine 6, 8 et 9 et qu'il reconnaît la constance que l'on retrouve dans la semaine 7.

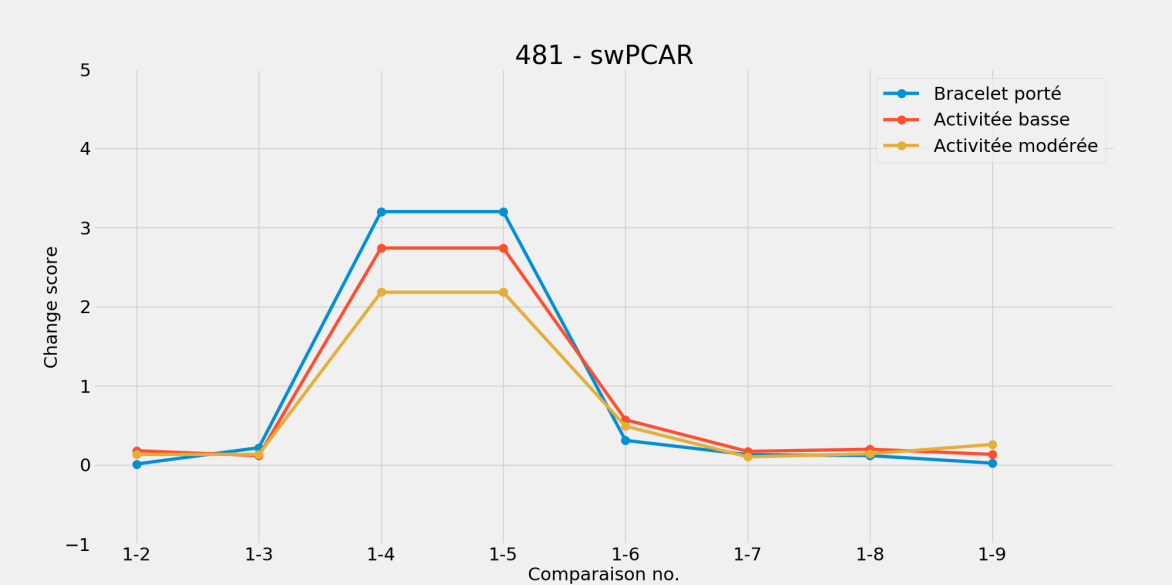


Figure 5.7 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme sw-PCAR

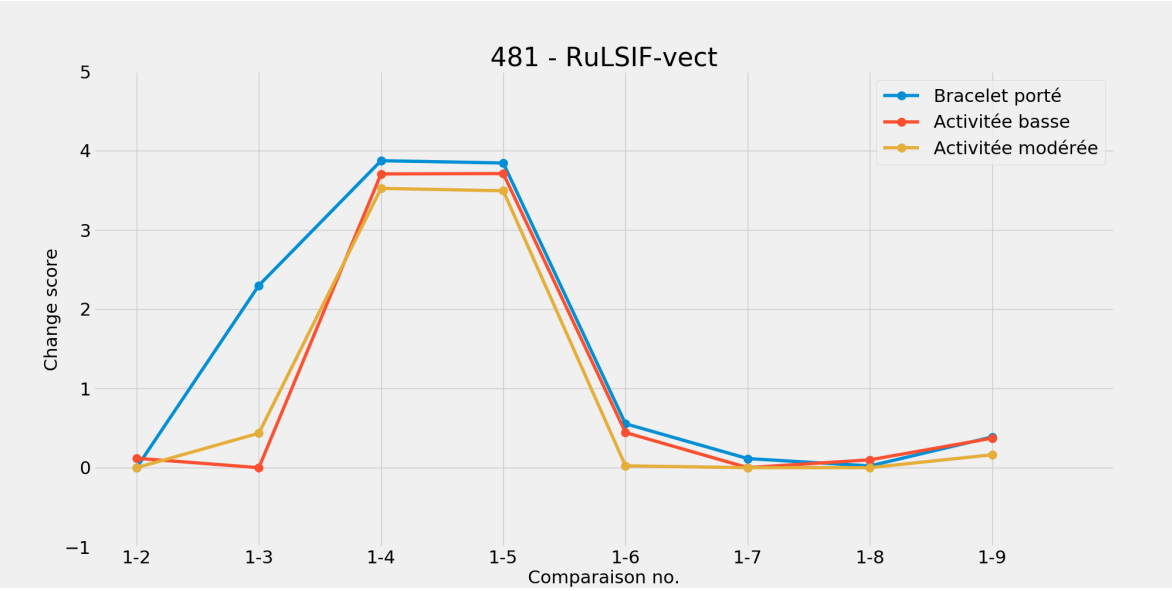


Figure 5.8 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode vecteur

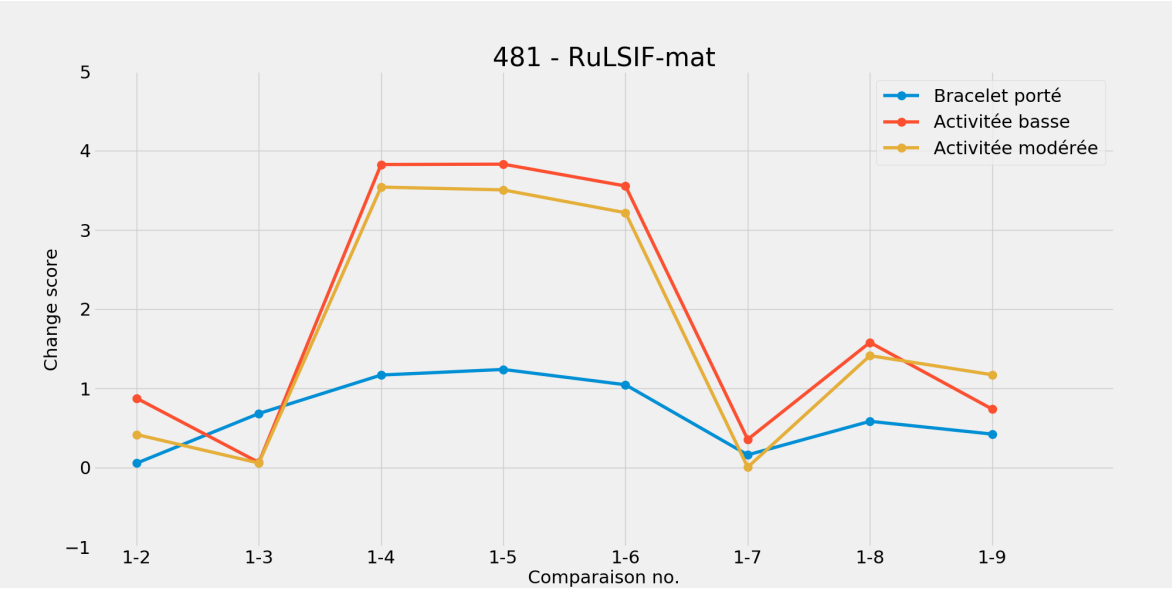


Figure 5.9 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice

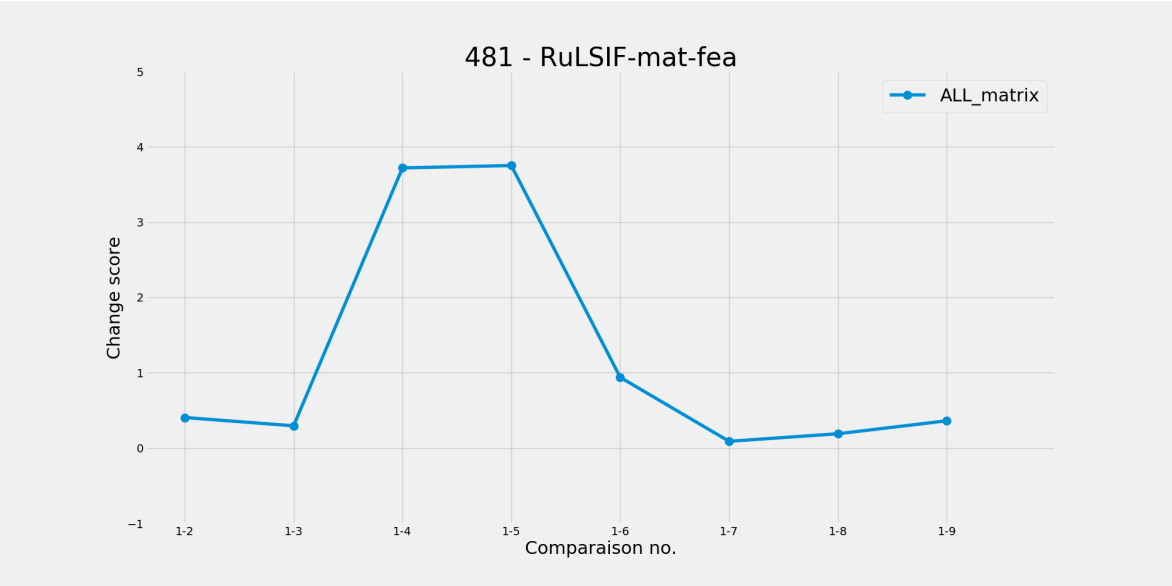


Figure 5.10 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice de caractéristiques

5.2.2 PARTICIPANT 907

Le participant 907 a débuté la cueillette en date du 3 octobre 2019 et neuf semaines complètes ont été enregistrées. Le carnet ne rapporte aucun événement de vie, mais il est à noter que ce participant est relativement actif. En effet, beaucoup de sorties du domicile sont rapportées.

Contrairement aux autres participants, beaucoup de données manquantes sont observées pour celui-ci. De plus, comme on peut le remarquer à l'aide des Figures 5.11 et 5.12, le temps où le bracelet était hors de portée de l'unité centrale de traitement (mode hors-ligne) surpasse le temps pour lequel le bracelet était connecté. Même si cette personne était très active, il

est très improbable que les déconnexions rapportées fussent toutes causées par des sorties du domicile.

Lors de l'installation des dispositifs de cueillette de données, il a été remarqué que le domicile du participant était une grande maison à deux étages (étage principal et sous-sol). L'unité centrale de traitement pouvait être installée seulement à l'étage principal à un endroit qui est près d'une extrémité de la maison. Le sous-sol était aménagé et le participant y passait un certain temps pour y effectuer des passe-temps. La théorie la plus plausible expliquant les données manquantes et les déconnexions fréquentes s'explique par la portée limitée de la technologie BLE. En effet, la portée limitée combinée avec un grand domicile de plusieurs étages peut causer des déconnexions fréquentes entre le bracelet et l'unité centrale de traitement. Même si la technologie développée fonctionne tout de même en cas de déconnexion, elle n'est pas conçue pour gérer des déconnexions trop fréquentes. Lorsqu'une déconnexion survient, un délai est observé avant que le bracelet ne détecte celle-ci, ce qui peut causer un nombre limité de données perdues à chaque déconnexion. Vraisemblablement, c'est ce qui s'est passé et c'est pourquoi le nombre de données manquantes est élevé de même que le temps hors connexion.

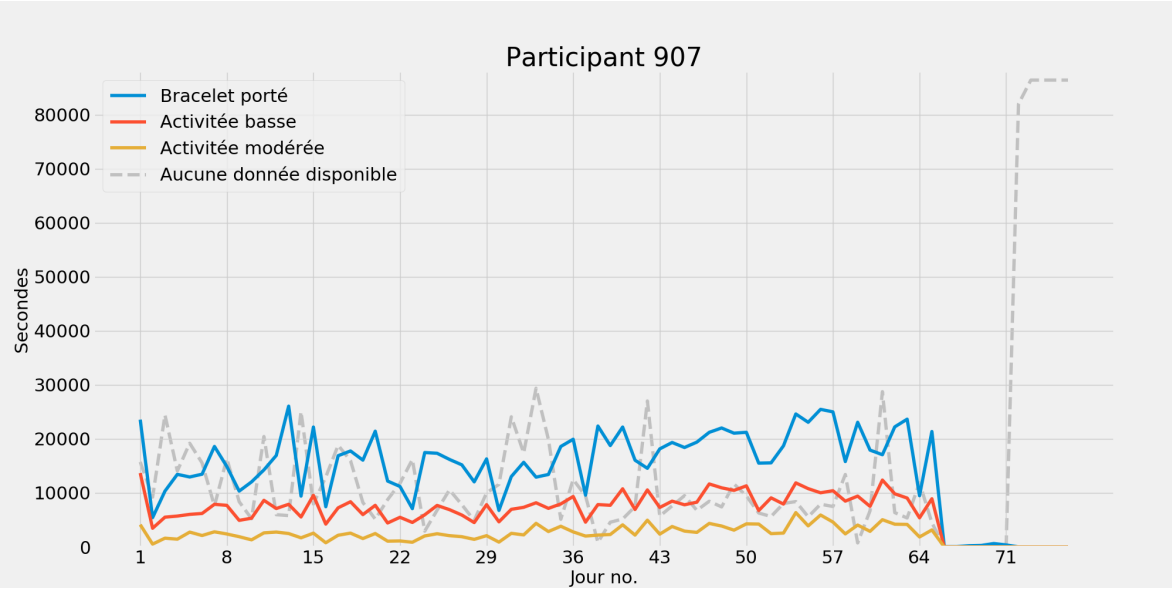


Figure 5.11 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 907

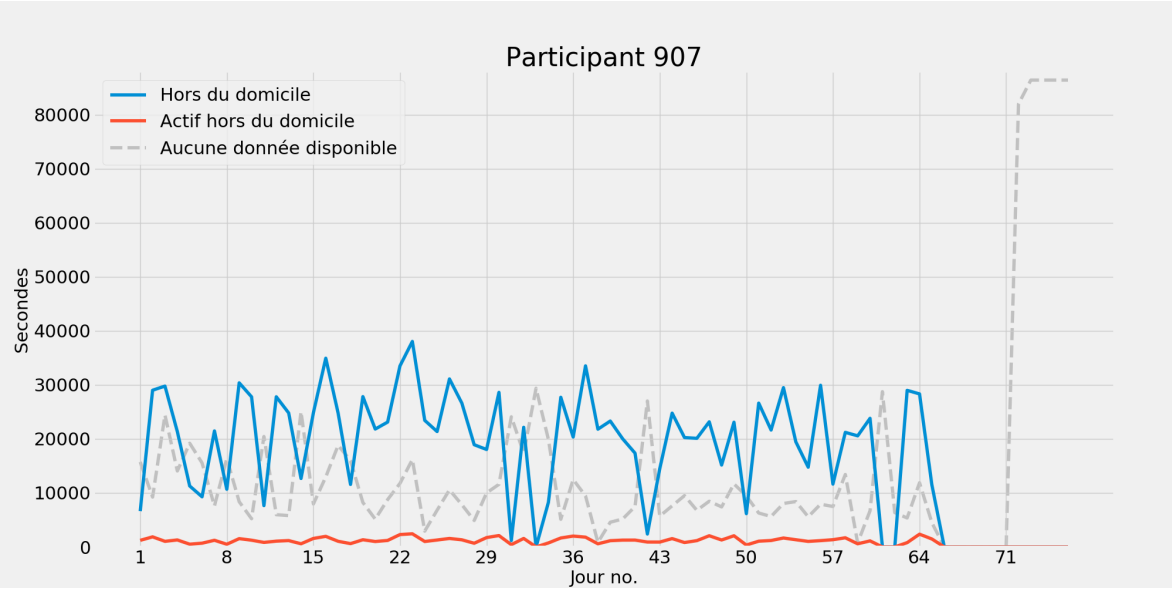


Figure 5.12 : Aperçu du temps écoulé lorsque le bracelet était hors de portée du domicile pour chaque catégorie d'action effectuée par le participant numéro 907

Détection de changement comportemental

Comme on peut le remarquer avec les Figures 5.13 et 5.16, sw-PCAR et le mode matrice de caractéristiques de RulSIF détectent des scores de changement plus importants dans les premières semaines. Cela n'est pas partagé par l'algorithme RulSIF-vecteur qui détecte très peu de changement. Finalement, le mode matrice de RulSIF rapporte des scores de changement élevés pour les activités modérées pour la plupart des semaines. Il est à noter que les données manquantes présentes pour ce participant sont trop importantes et le résultat des algorithmes de détection de changement comportemental trop divergent pour tirer des conclusions.

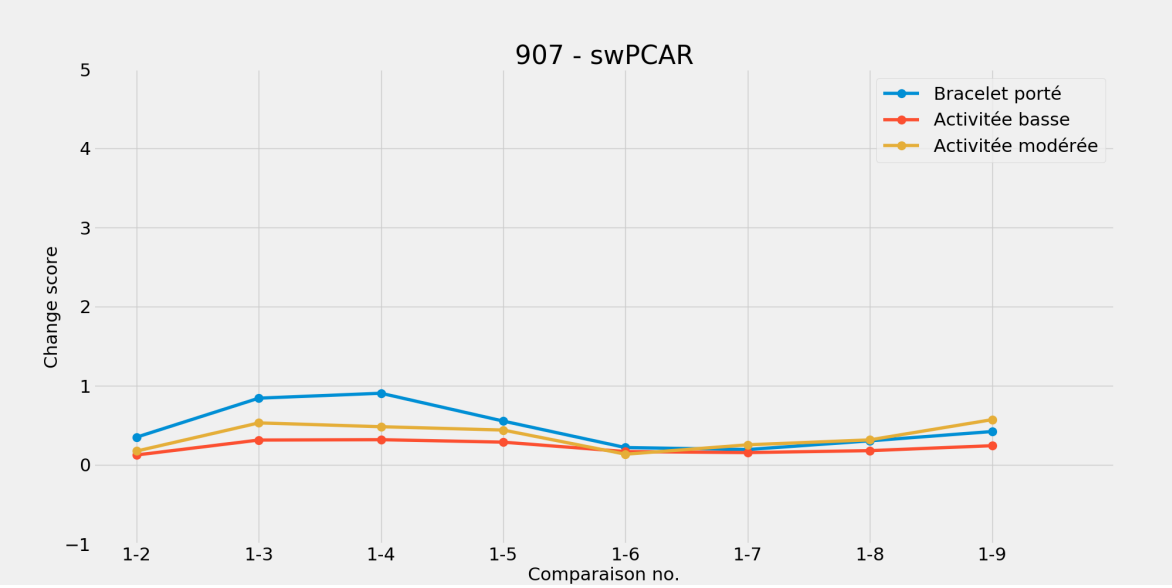


Figure 5.13 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme sw-PCAR

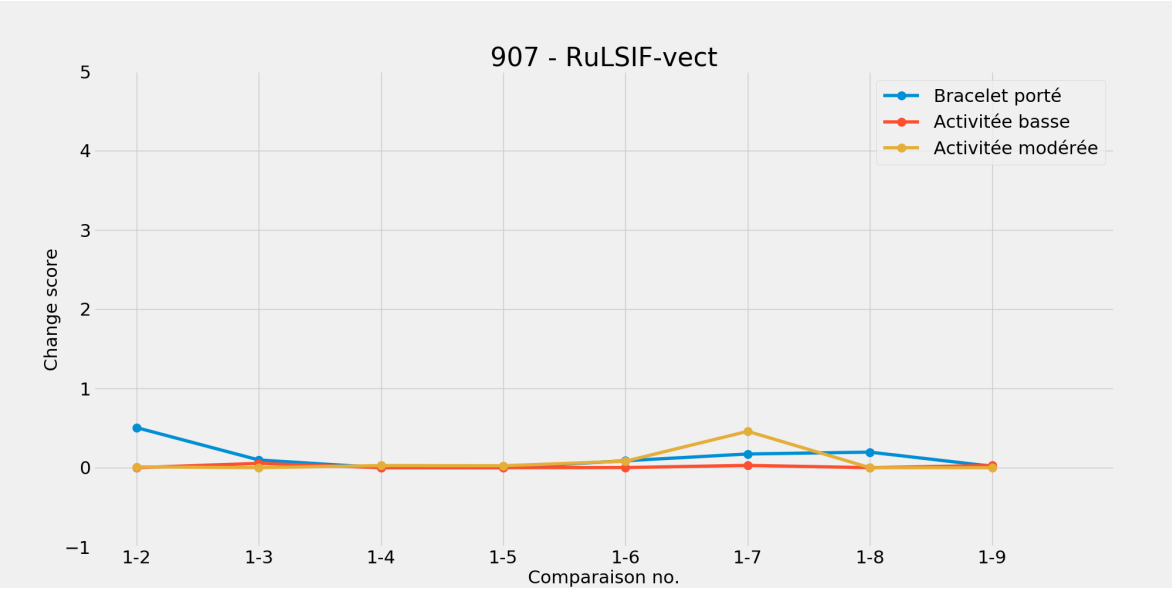


Figure 5.14 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode vecteur

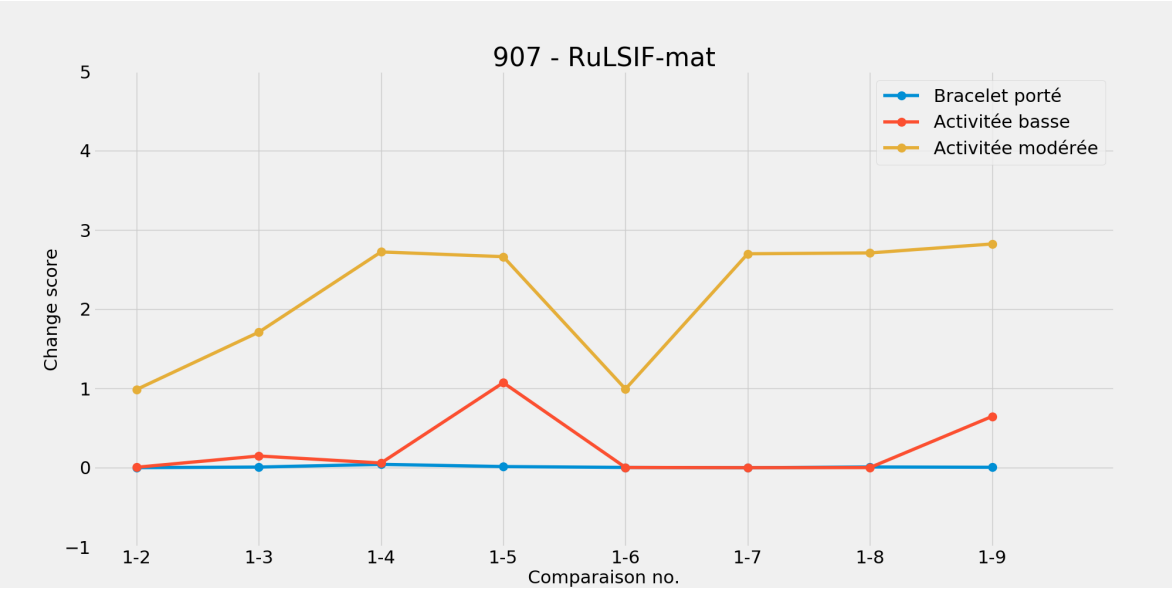


Figure 5.15 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice

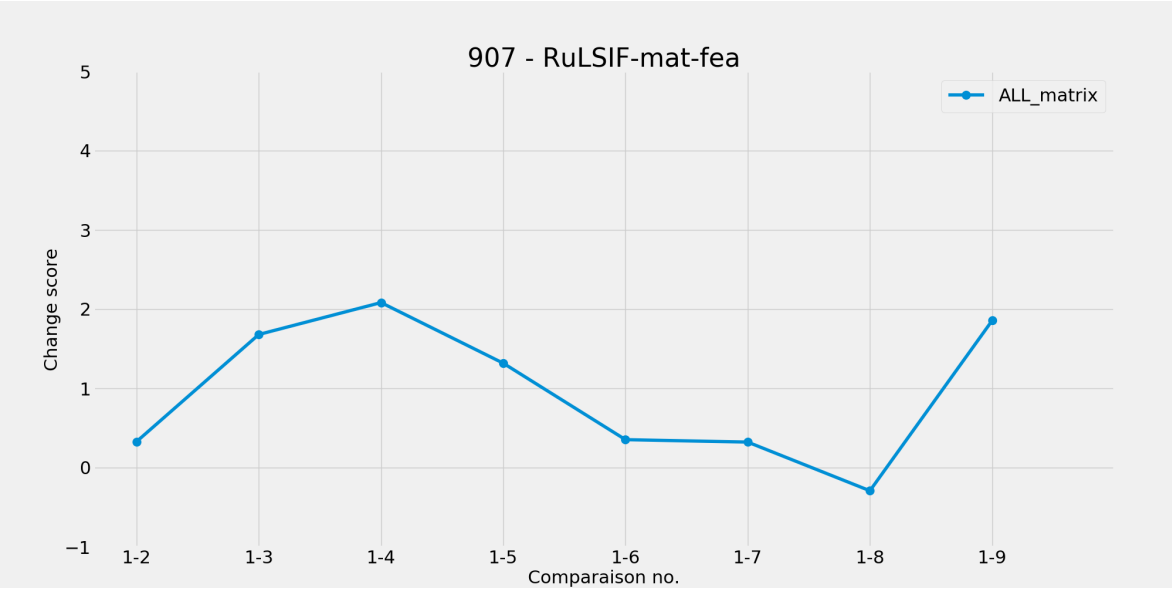


Figure 5.16 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice de caractéristiques

5.2.3 PARTICIPANT 1453

Le participant 1453 a commencé l'expérimentation en date du 5 octobre 2019 et seulement quatre semaines de données sont disponibles. En effet, la carte mémoire de la centrale s'est corrompue pendant l'expérimentation et seulement les quatre premières semaines ont pu être récupérées à l'aide des sauvegardes. Aussi, aucun événement de vie n'est inscrit dans le carnet de bord.

À titre informatif, la Figure 5.17 contient un graphique qui illustre le temps écoulé pour chaque catégorie d'action effectuée. On voit facilement les quatre semaines où des données ont pu être récupérées. On observe 2 journées où le bracelet n'a pas été porté et un peu plus d'une

journée où le bracelet n'était pas opérationnel. La théorie la plus plausible pour expliquer cette journée non opérationnelle s'explique par le déchargement complet de la batterie du bracelet qui a été rechargée une journée plus tard. On peut aussi remarquer que la dernière semaine n'est pas complète, ce qui laisse seulement trois semaines pour les analyses de changement comportemental. Ce nombre est insuffisant pour que les résultats de l'analyse de changement comportemental soient pertinents, c'est pourquoi elle ne sera pas présentée.

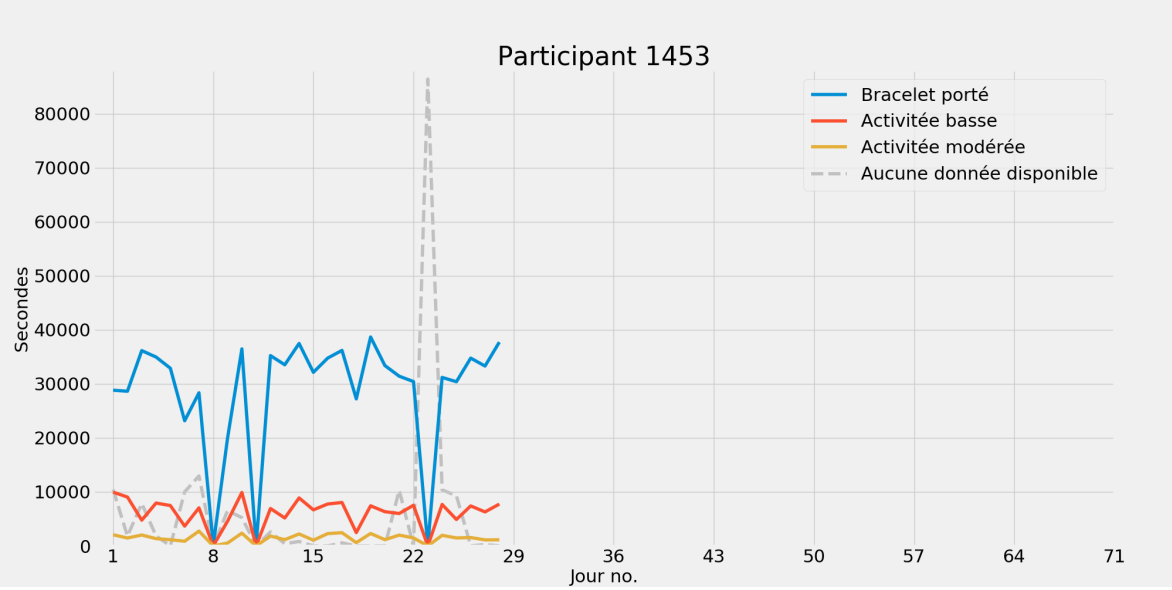


Figure 5.17 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 1453

5.2.4 PARTICIPANT 1692

Le participant 1692 a débuté la cueillette de données en date du 4 octobre 2019 et les données de neuf semaines complètes ont pu être recueillies. Le carnet du participant ne

mentionne aucun événement de vie qui aurait pu affecter ses habitudes et aucun bris matériel n'a été observé.

La Figure 5.18 illustre le temps écoulé pour chaque catégorie d'action effectuée. On remarque que la routine de l'individu est quelque peu instable et que certaines journées comportent un non-port du bracelet. Néanmoins, c'est un bon candidat pour l'application des algorithmes de détection de changement comportemental.

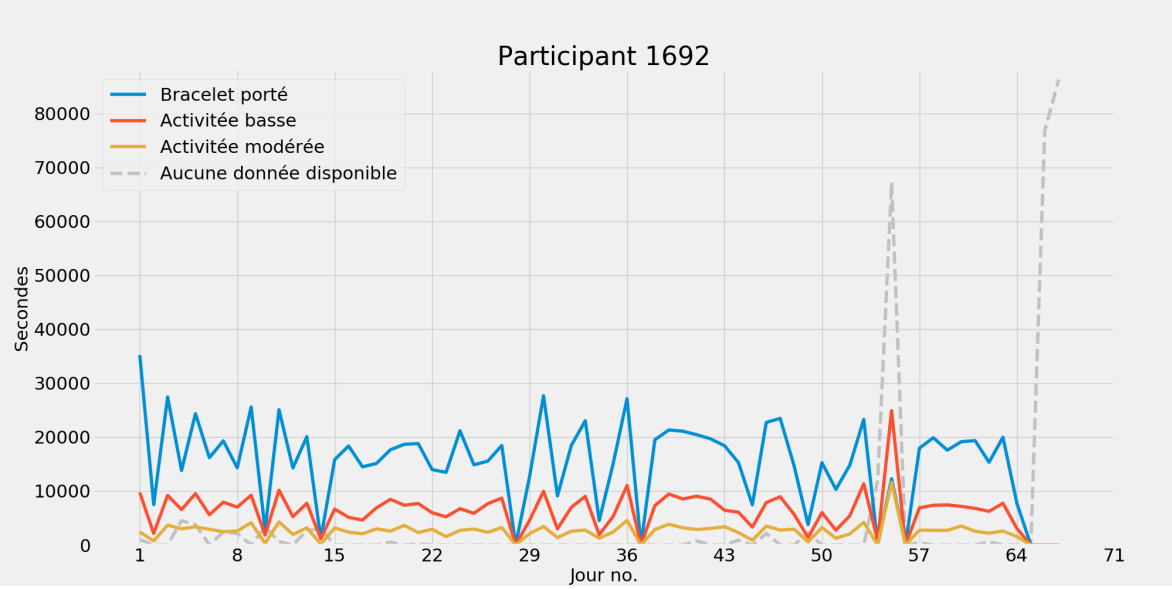


Figure 5.18 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 1692

Détection de changement comportemental

Les irrégularités observées à l'aide de la Figure 5.18 affectent de façon considérable les scores de changement. Contrairement aux autres algorithmes, sw-PCAR génère des scores

de changement assez stables entre les différentes semaines. On peut cependant observer un léger pic pour les scores de la semaine 4 et la semaine 7. Les modes d'exécution vecteur et matrice de RulSIF sont beaucoup moins stables que sw-PCAR et les scores de changement sont relativement élevés. Pour le mode d'exécution vecteur de RulSIF, on observe la plus grande variation de score de changement à la semaine 7 et les activités basses sont celles qui génèrent les scores de changement les plus élevés. Le mode matrice de RulSIF est largement affecté par les irrégularités dans la routine et les scores de changement associés sont à la fois élevés et instables. Finalement, le dernier mode d'exécution de RulSIF est beaucoup plus stable et un pic mineur à la semaine 4 est observé tout comme sw-PCAR.

Même si aucun événement de vie n'a été rapporté dans le carnet de bord de ce participant, trois algorithmes sur quatre ont détecté un changement de comportement plus élevé lors de la semaine 4. On peut donc supposer qu'un léger changement de comportement a eu lieu lors de cette semaine, mais il est impossible de le vérifier avec les informations dont nous disposons. On a aussi pu remarquer des irrégularités dans la routine de l'individu et les algorithmes ont réagi à cela de façon marquée. Le mode d'exécution matrice de RulSIF est celui qui a été le plus marqué par ces irrégularités.

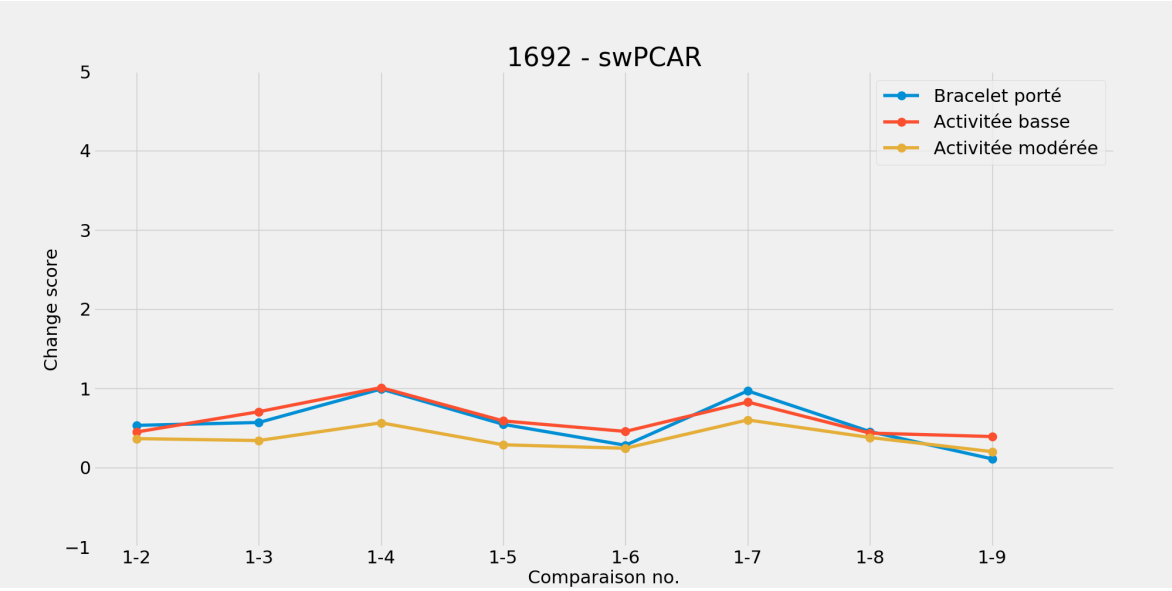


Figure 5.19 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme sw-PCAR

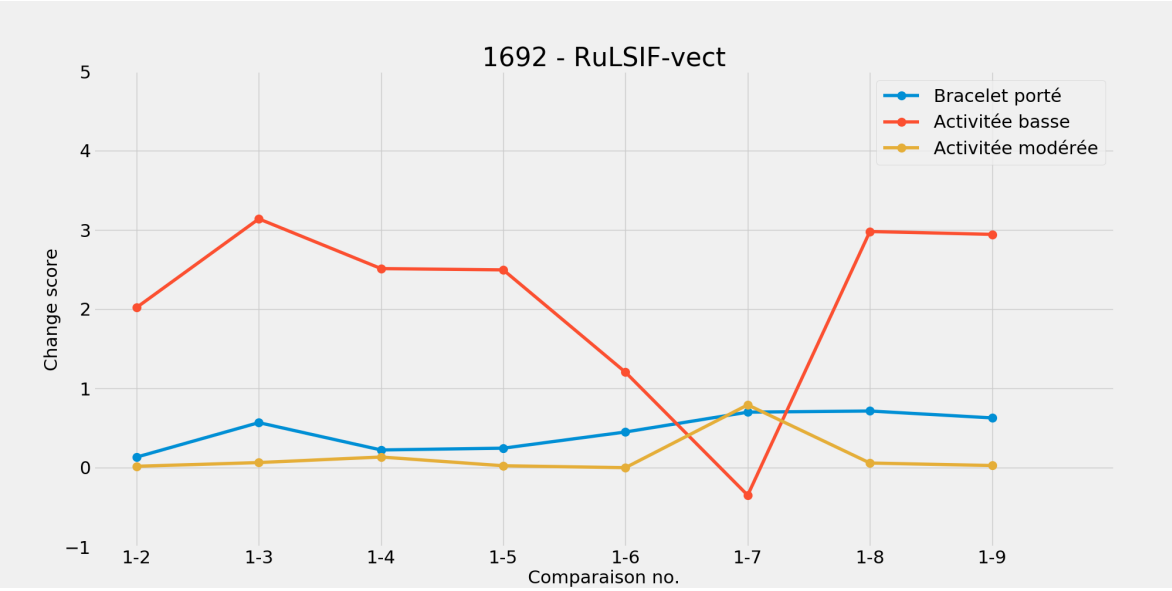


Figure 5.20 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode vecteur

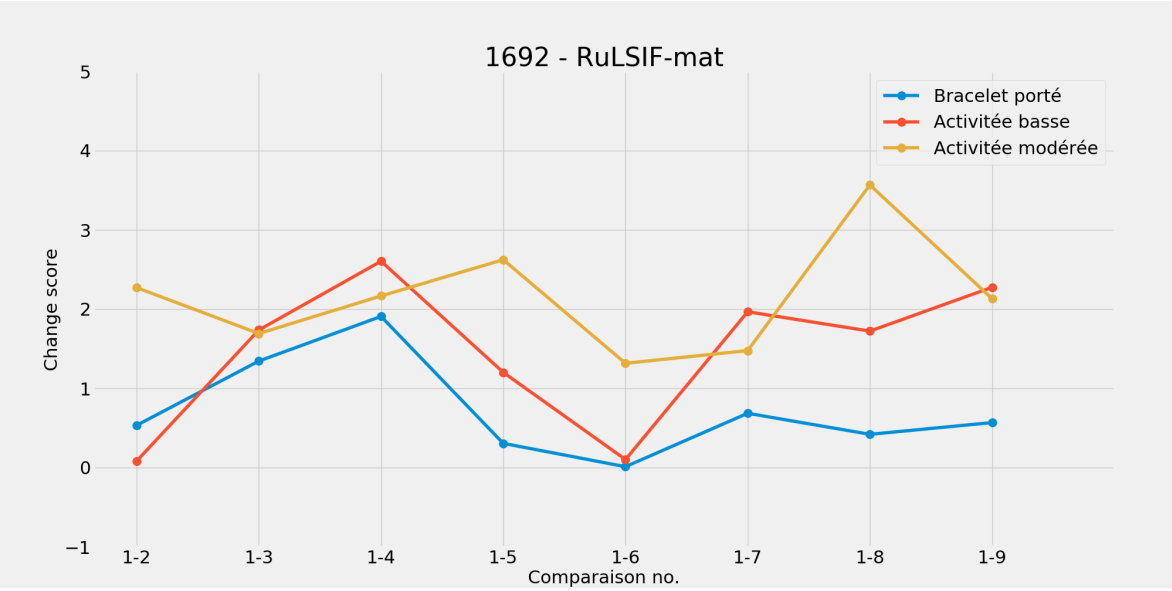


Figure 5.21 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice

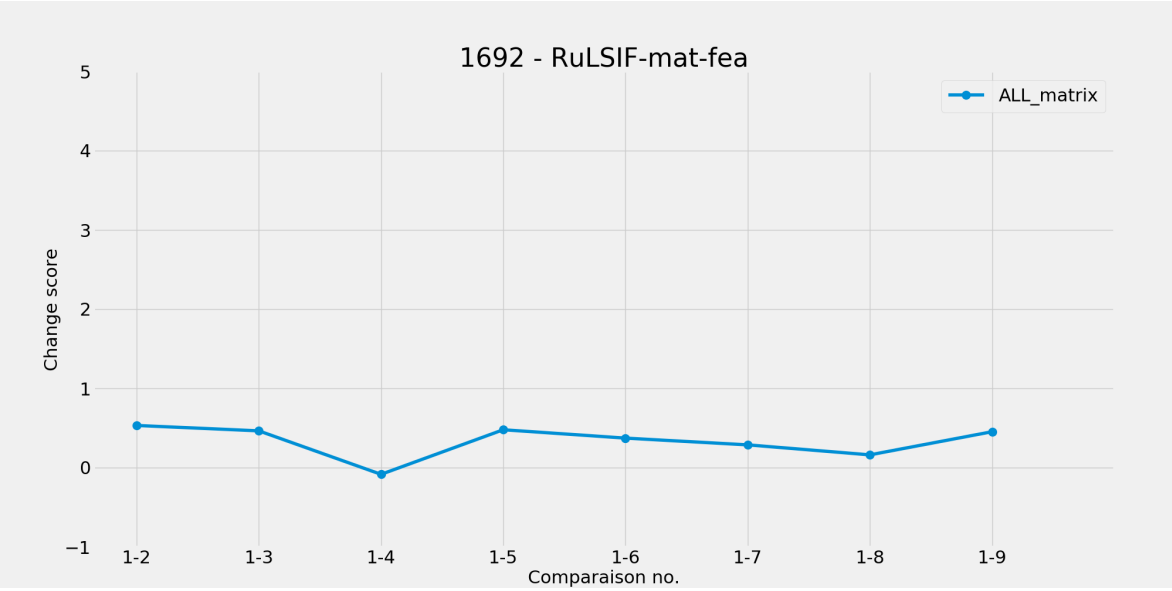


Figure 5.22 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice de caractéristiques

5.2.5 PARTICIPANT 2002

La cueillette de données du participant 2002 a débuté le 2 octobre 2019 et un total de neuf semaines complètes sont disponibles. Le carnet de bord signale deux événements de vie qui pourraient affecter la routine de l'individu :

— Une gastro-entérite a été rapportée lors de la quatrième semaine.

— L'influenza a été rapporté lors de la neuvième semaine.

La Figure 5.23 contient un graphique qui illustre le temps écoulé pour chaque catégorie d'action effectuée. Le port du bracelet a été généralement respecté. On peut voir que le bracelet n'a pas été porté pendant presque trois jours lors de la neuvième semaine et pendant 2 jours lors de la quatrième semaine. Malheureusement, ces semaines correspondent aux deux semaines où des événements de vie ont été rapportés. On remarque aussi que le niveau de port du bracelet varie considérablement de jour en jour.

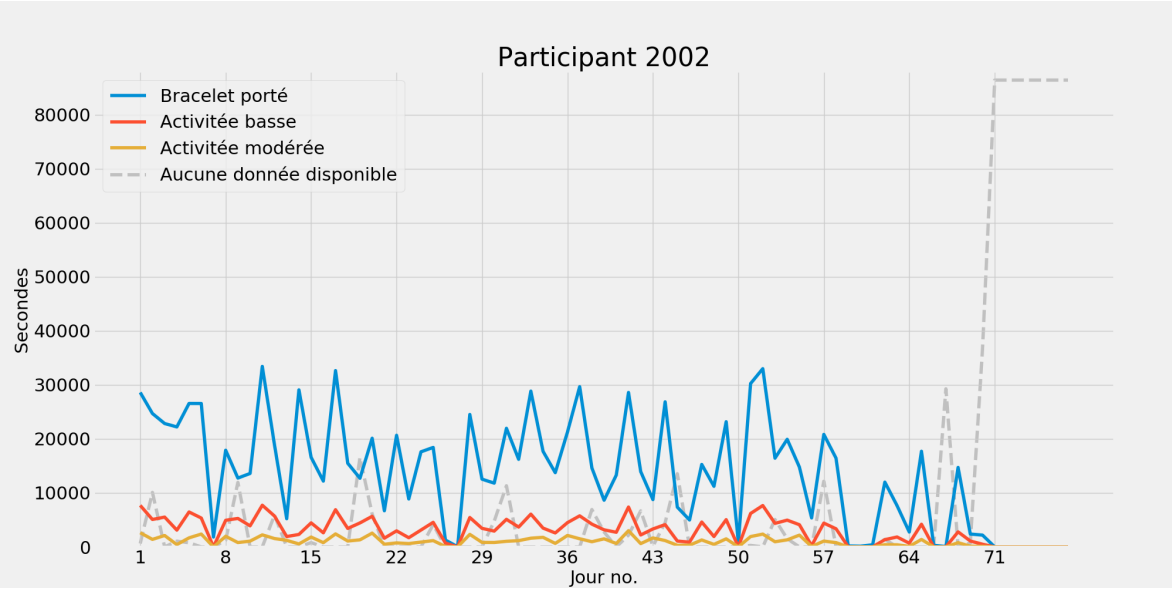


Figure 5.23 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 2002

Détection de changement comportemental

Comme le bracelet ne semble pas avoir été porté lorsque le participant était malade, il est difficile d'évaluer l'impact des deux événements de vie dans la routine de l'individu. En effet, les expérimentations ont démontré qu'une modification de la routine (ex. : changement des heures où des activités sont effectuées) affecte plus sérieusement les scores de changement qu'une journée sans données. De ce fait, il sera difficile de démontrer un quelconque changement de routine lié à ces événements. Remarquons tout de même une légère augmentation du score de changement pour la neuvième semaine (semaine où l'influenza a été rapporté) émis par l'algorithme RulSIF-vect. On remarque aussi un léger pic lors de la semaine 8 pour les trois autres algorithmes. De plus, un léger pic est observé lors de la quatrième semaine

(semaine où une gastro-entérite a été rapportée) pour l'algorithme RulSIF-matrice. Cependant, ces pics ne sont pas assez prononcés pour en tirer des conclusions.

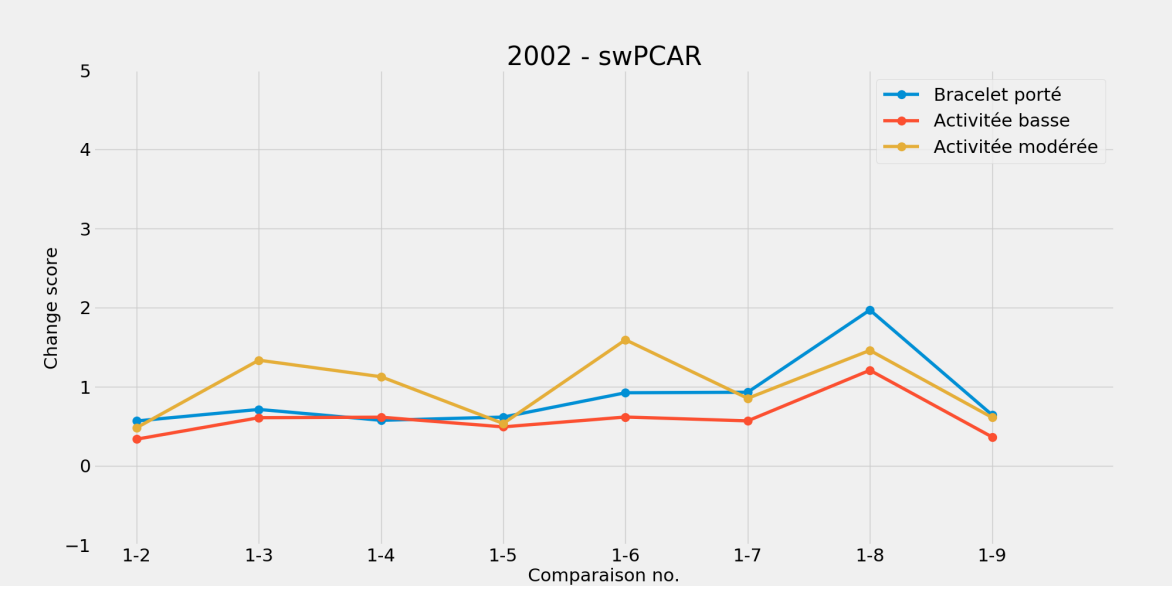


Figure 5.24 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme sw-PCAR

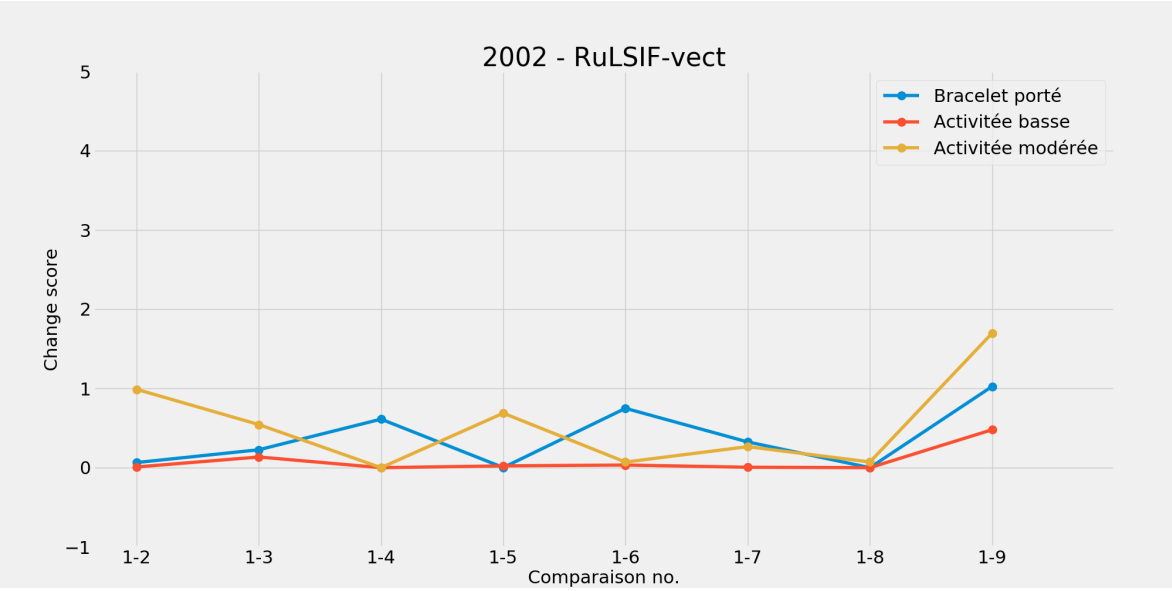


Figure 5.25 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode vecteur

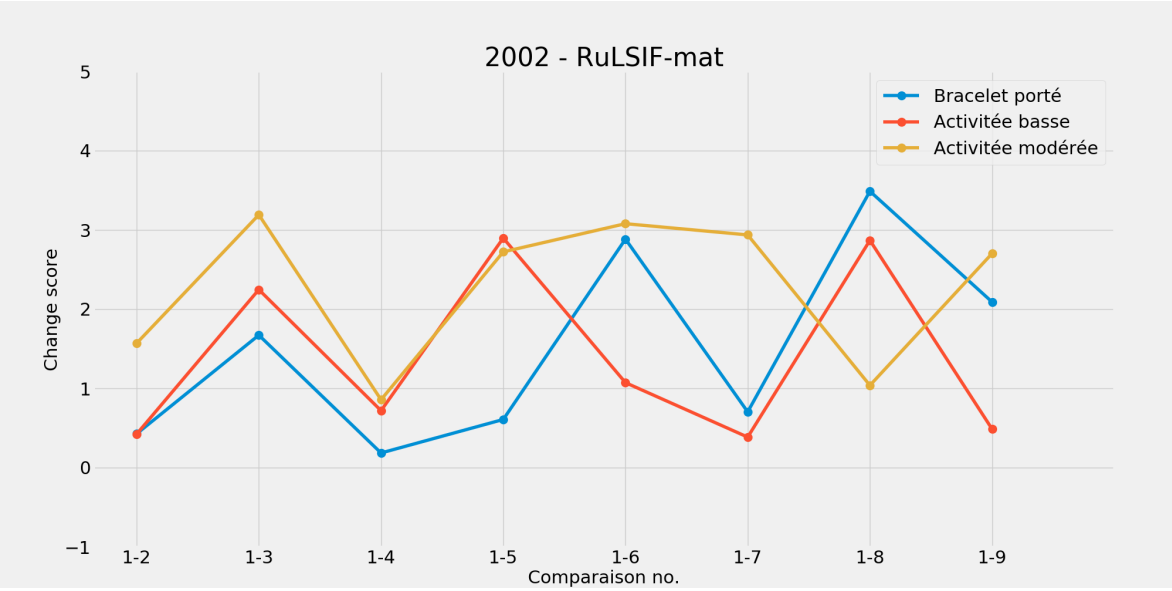


Figure 5.26 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice

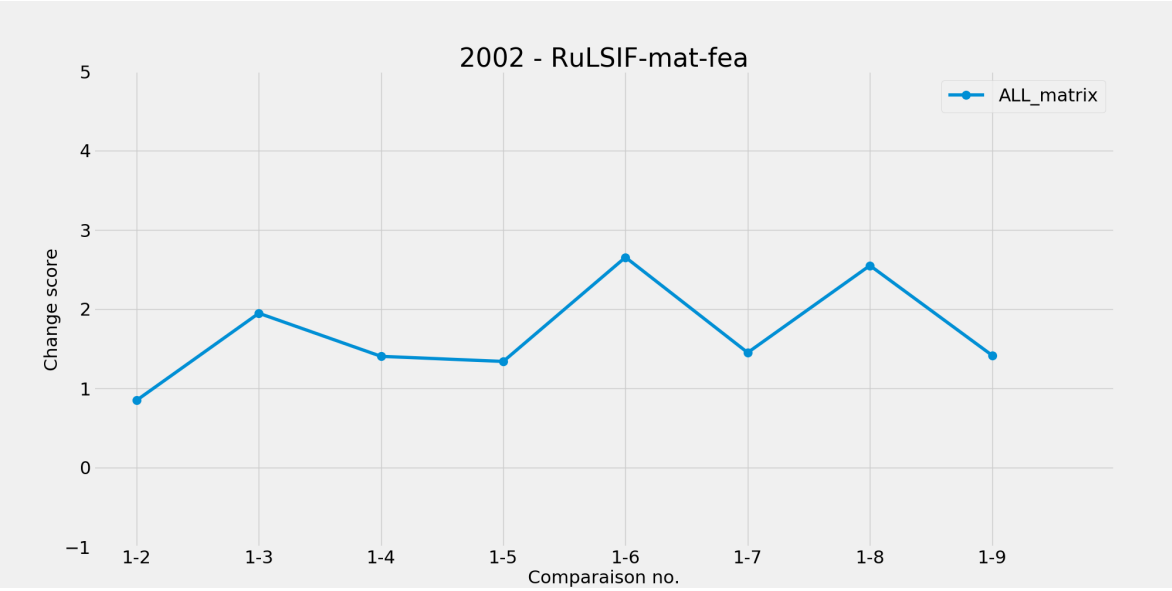


Figure 5.27 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice de caractéristiques

5.2.6 PARTICIPANT 2182

Le participant numéro 2182 a débuté la cueillette en date du 3 octobre 2019 et les données de huit semaines complètes ont pu être recueillies. Même si celui-ci fait partie des cinq personnes ayant arrêté la cueillette avant la fin, huit semaines complètes sont tout de même disponibles. Le carnet du participant ne renferme aucune information qui pourrait indiquer un événement de vie.

Le port du bracelet pour ce participant est irrégulier. Cependant, à l'aide de la Figure 5.28, il est possible de remarquer une légère tendance traduisant une diminution du port du bracelet avec le temps. On observe aussi deux journées pour lesquelles le bracelet n'était

pas opérationnel, cela peut être expliqué par les facteurs mentionnés précédemment tel un déchargement complet de celui-ci.

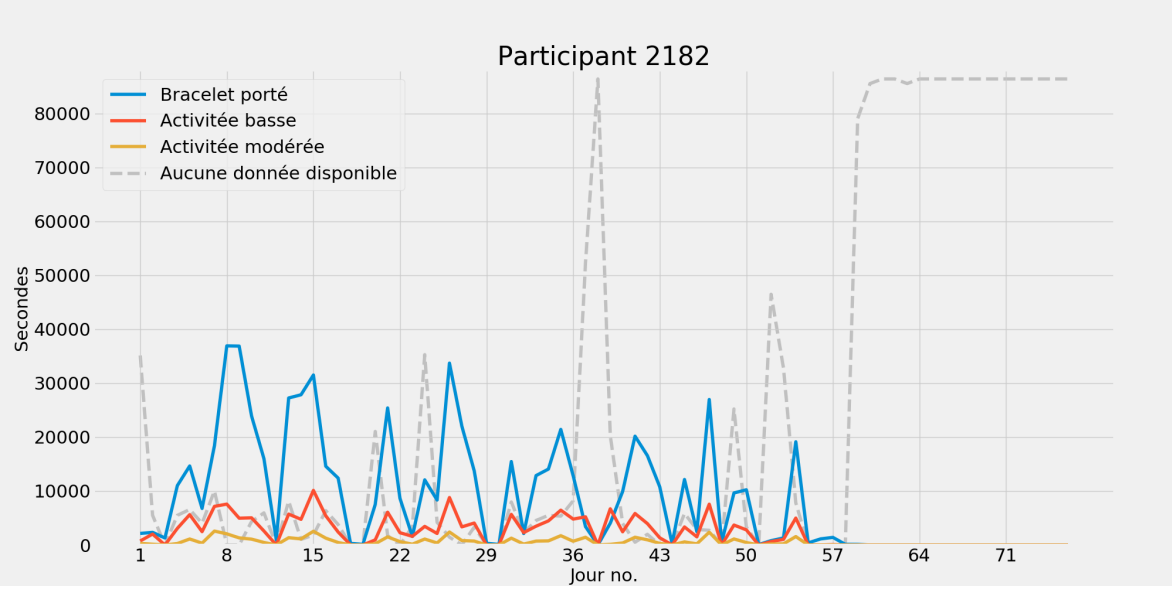


Figure 5.28 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 2182

Détection de changement comportemental

Tous les algorithmes à l'exception du mode d'exécution vecteur de RulSIF traduisent la tendance observée précédemment. En effet, les scores de changement augmentent légèrement chaque semaine. Bien que le mode d'exécution matrice de RulSIF indique des changements principalement au niveau des activités faibles, on ne peut tirer des conclusions quant à l'augmentation ou la diminution de l'activité physique au cours des huit semaines analysées. En effet, trop d'instabilités quant au port du bracelet sont observées, ce qui peut introduire un

biais.

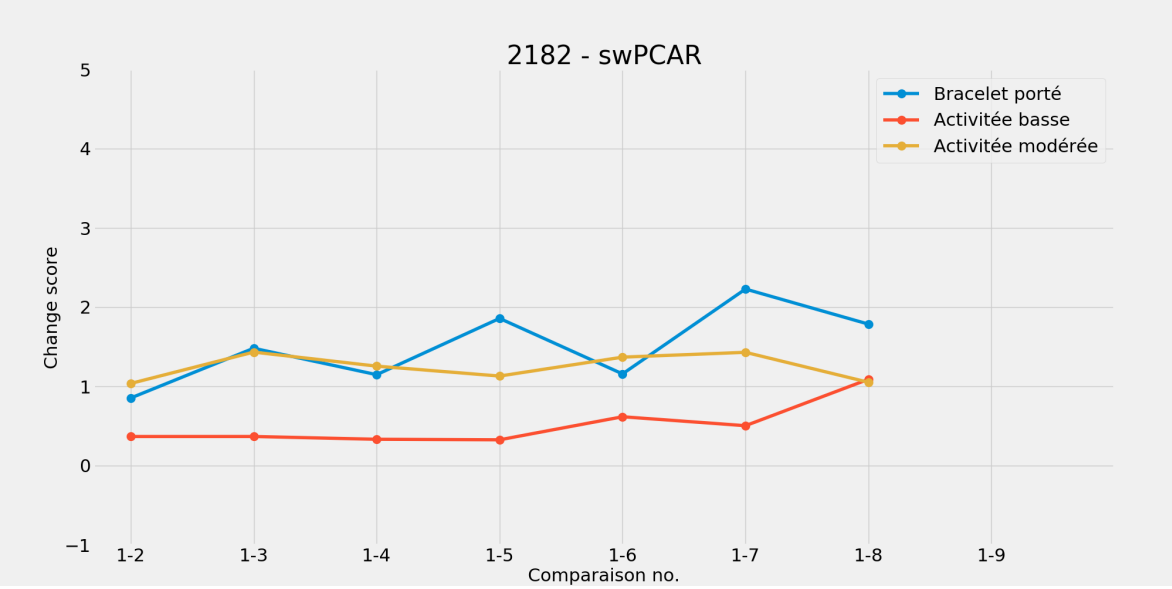


Figure 5.29 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme sw-PCAR

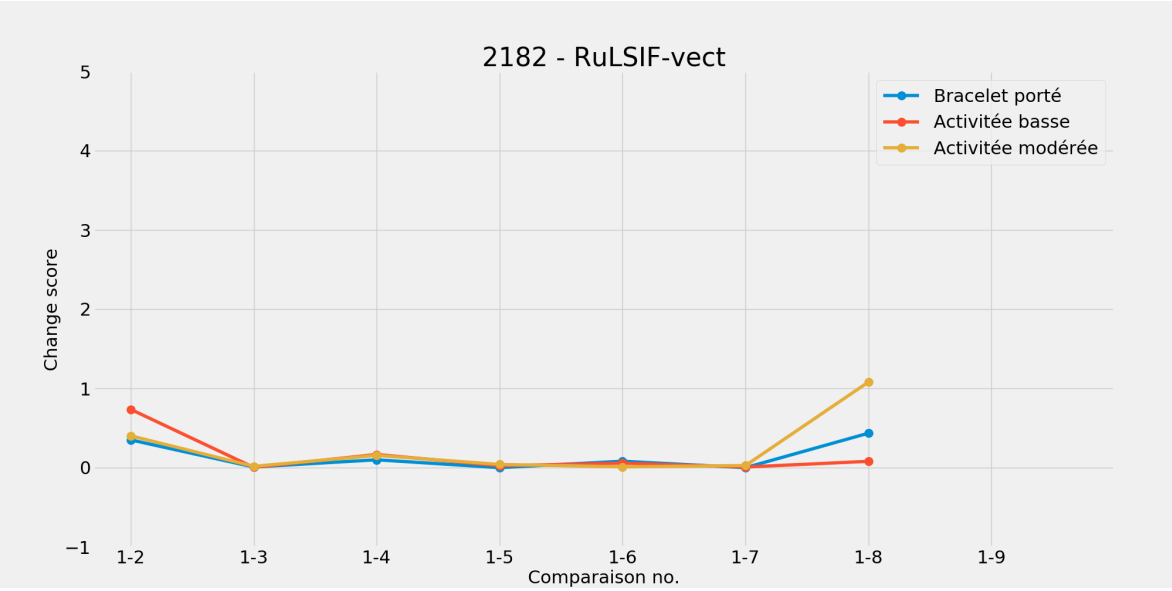


Figure 5.30 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode vecteur

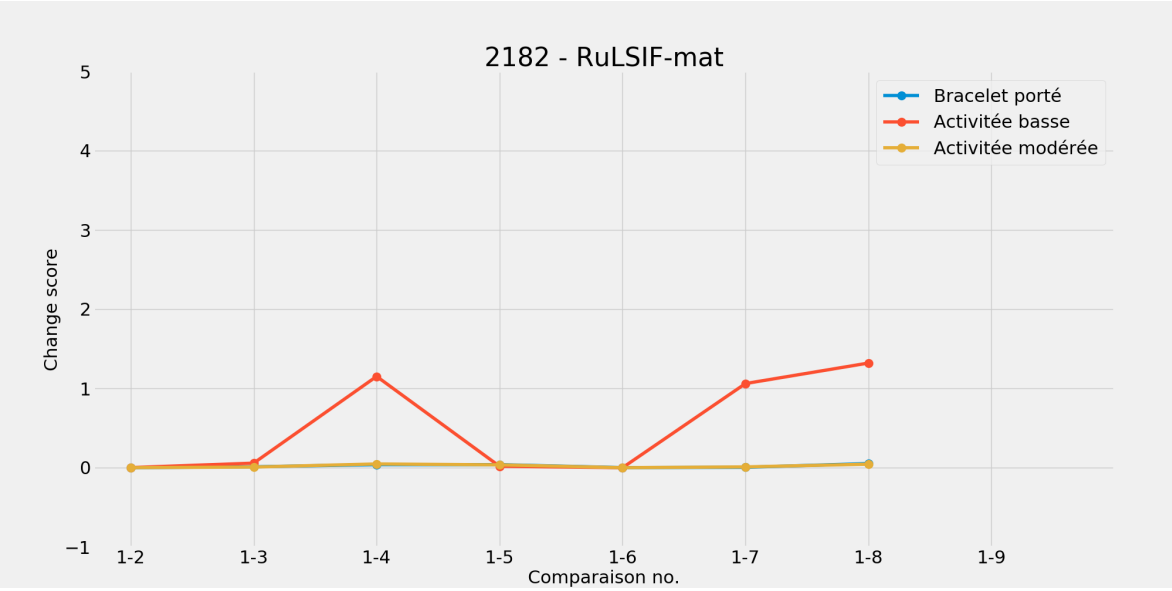


Figure 5.31 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice

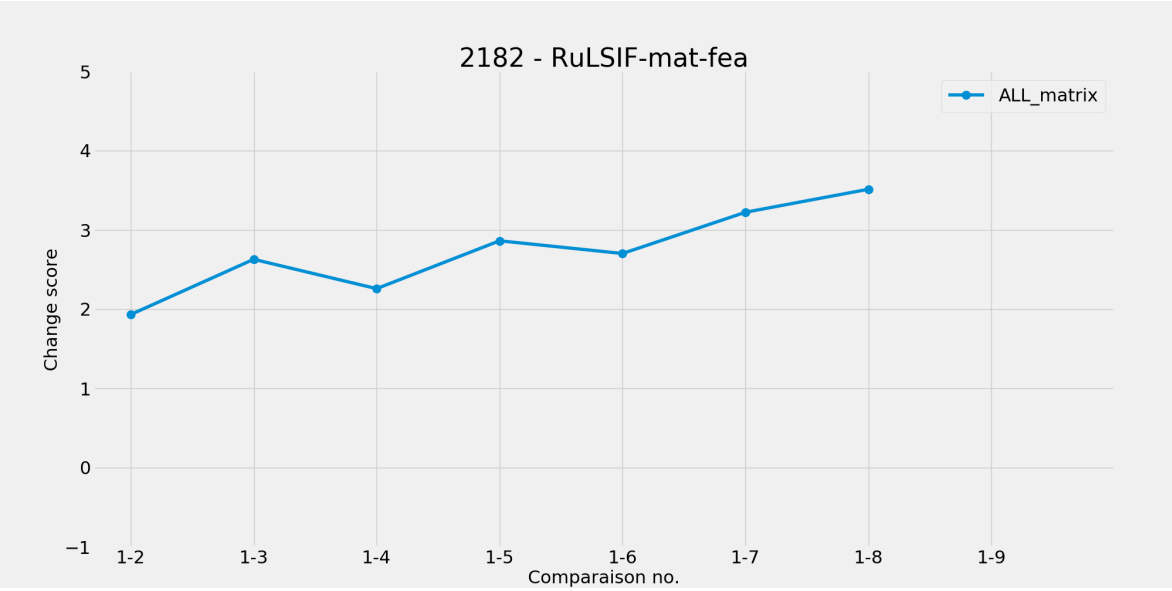


Figure 5.32 : Graphique illustrant les scores de changement obtenus à l'aide de l'algorithme RulSIF en mode matrice de caractéristiques

5.3 RÉSULTATS GÉNÉRAUX

La cueillette de données a permis de recueillir des informations sur les activités effectuées chaque jour pour une durée de dix semaines. Un total de dix personnes étaient initialement inscrites pour cette cueillette, mais six d'entre elles ont participé pendant au moins huit semaines. Comme mentionné plus tôt, les autres ont dû se désister en cours de route pour des raisons non liées à la technologie. La section précédente s'est occupée d'effectuer l'analyse des changements comportementaux de chacun de ces individus. Pour cette section, il sera question d'analyser les résultats de façon plus générale. En effet, la participation générale des individus ainsi que la constance de leur routine seront évaluées et les événements de vie seront explorés plus en détail.

5.3.1 PARTICIPATION

L'étude effectuée dans la cadre du projet Acti-DM1 comportait à la base 20 personnes. La moitié faisait partie du groupe ayant l'aide de la technologie et l'autre moitié constituait le groupe de contrôle qui n'avait pas l'aide de la technologie pour la réalisation de leur programme d'entraînement. Le taux d'abandon du groupe de contrôle est de 10% (une personne a quitté) et le taux d'abandon du groupe technologique est de 40% (quatre personnes ont quitté). Le taux d'abandon global est donc de 25%. Comme le groupe de contrôle n'avait pas l'aide de la technologie, il ne sera pas possible d'évaluer leur participation davantage. Pour la suite, seulement les six personnes restantes du groupe technologique seront considérées.

Le Tableau 5.1 présente les statistiques liées au port du bracelet lors de la cueillette. Le nombre de jours analysés correspond aux jours considérés pour ce participant et est fidèle au nombre de jours considérés par l'étude des changements comportementaux. Le nombre de jours valides inclut les journées comportant un nombre minimal de données manquantes. Lorsque le nombre de données manquantes pour une journée excède trois heures, la journée est considérée comme invalide. On voit bien ce phénomène avec le participant numéro 481 pour lequel le bracelet a été en panne pour un peu plus de 2 semaines. Sur les 63 jours analysés, 45 sont valides, car les données de près de 18 jours sont manquantes. Le nombre de jours pour lesquels le bracelet a été porté correspond aux jours valides où le bracelet a été porté au moins quatre heures. Le taux de port est calculé à partir du nombre de jours valides et du nombre de jours où le bracelet a été porté. Finalement, la moyenne du temps où le bracelet a été porté en heures est calculée à partir de temps de port enregistrés lors des journées valides.

Tableau 5.1 : Participation des participants à la cueillette

Participant	Nb. jours Analysés	Nb. jours valides	Nb. jours bracelet porté	% porté	Moyenne du port par jour (h)
481	63	45	38	84%	7,14 h
907	63	39	33	85%	5,09 h
1453	29	26	24	92%	8,32 h
1692	63	60	43	72%	4,55 h
2002	63	57	34	60%	4,57 h
2182	56	46	19	41%	3,65 h

La participation varie considérablement selon le participant. Certains ont participé de manière assidue (ex. : participant 1453 avec 92% de port et 8,32h par jour de moyenne) alors que certains ont participé beaucoup moins minutieusement (ex. : participant 2182 avec 41% de port et 3,65h par jour de moyenne). En somme, le taux global moyen de port du bracelet est de 70% et le nombre d'heures moyen de port par jour est de 5,02h.

5.3.2 CONSTANCE DE LA ROUTINE

Lors de la cueillette de données, les participants devaient porter leur bracelet chaque jour afin d'enregistrer des informations sur leur niveau d'activité. Une personne peut porter son bracelet selon une certaine logique (ex. : début et fin du port à un certain moment de la journée) et effectuer certaines activités durant le port de son bracelet qui va varier le niveau d'activité enregistré. Afin d'évaluer à quel point la routine de chaque participant était stable durant la cueillette, un algorithme utilisant sw-PCAR a été mis en place.

Le fonctionnement de l'algorithme est plutôt simple : celui-ci évalue à quel point chaque

semaine de la cueillette est différente des autres à l'aide des scores de changement rapportés par sw-PCAR. Comme il a été discuté plus tôt, la détection de changements comportementaux appliquée dans ce mémoire est configurée afin de comparer la première semaine avec toutes les autres, ce qui permet d'évaluer l'évolution du score de changement dans le temps selon cette semaine de référence. Le processus d'évaluation de la constance de la routine applique ce même principe, mais celui-ci est effectué autant de fois qu'il y a de semaines. En effet, ce processus débute en posant la première semaine comme la semaine de référence, puis évalue les scores de changement en la comparant avec toutes les autres. Ensuite, la moyenne des scores de changement obtenus avec cette semaine de référence est calculée puis sauvegardée. Un nouveau cycle peut alors commencer : la semaine de référence devient la deuxième semaine, des scores de changement sont calculés en la comparant avec toutes les autres et une moyenne de ces scores de changement est calculée. Ce principe s'effectue pour chaque semaine et va donner un score de changement moyen par semaine pour chaque caractéristique. Plus la moyenne des scores de changement d'une semaine est élevée, plus elle est différente des autres. Lorsque l'on veut généraliser encore plus, il est possible d'avoir une idée générale de la constance de la routine pour un participant en effectuant la moyenne des scores de changement moyen obtenus à la suite de l'application de ce processus. C'est d'ailleurs ce que la Figure 5.33 illustre. Le graphique illustré dans cette figure évalue la constance de la routine selon les scores de changement moyen globaux. Plus le score de changement d'une caractéristique est élevé, moins la routine est stable.

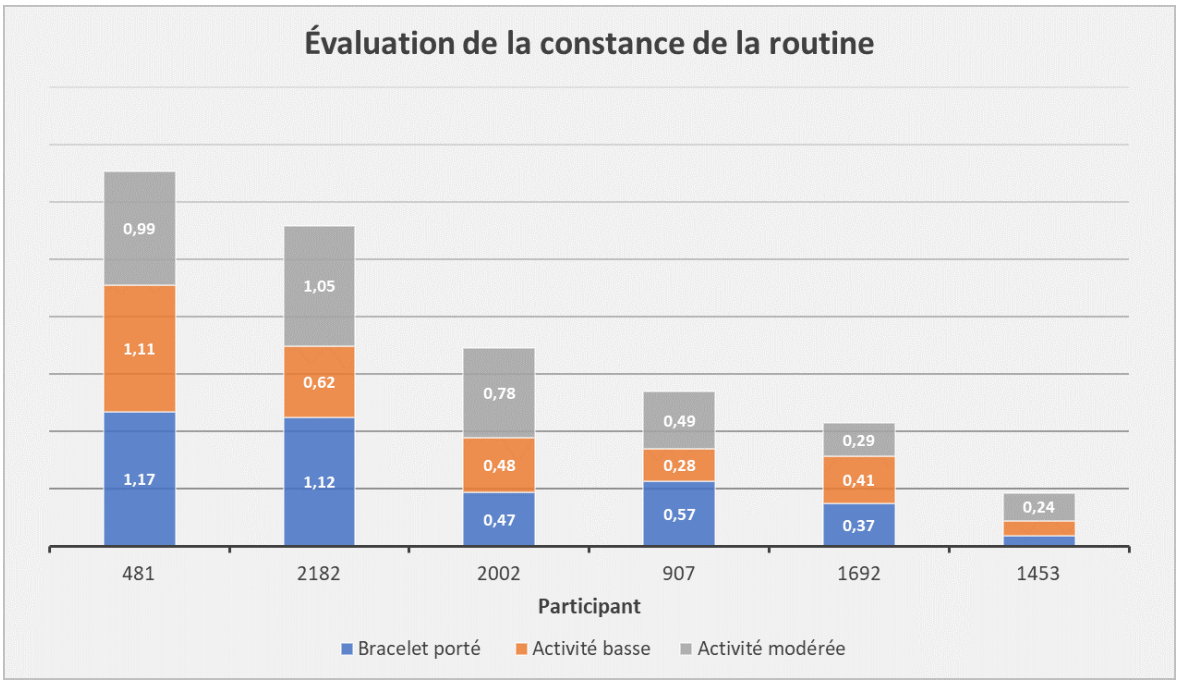


Figure 5.33 : Évaluation de la constance de la routine pour chaque participant

Tout comme pour la participation à l'expérimentation, la constance de la routine varie considérablement d'un individu à l'autre. La routine du participant 1453 est très stable, mais il est important de retenir que celui-ci comporte seulement 4 semaines de données. Il aurait été intéressant d'évaluer la constance de sa routine pour toute la durée du programme d'entraînement. Le participant 481 comporte une routine très inconstante, ce qui peut s'expliquer selon deux facteurs. Premièrement, deux semaines sans données sont présentes à la suite d'un bris du bracelet, ce qui affecte inévitablement la routine de ces semaines. Deuxièmement, un port du bracelet plus sporadique est observé à la suite du rétablissement de son bracelet (voir Figure 5.6), ce qui affecte d'autant plus sa routine.

5.3.3 ÉVÉNEMENTS DE VIE

Quatre événements de vie ont été signalés durant la cueillette de données pour deux personnes appartenant au groupe technologique. Les deux premiers événements se rapportent à une personne qui a quitté en cours de route pour des raisons de chirurgie. Cette personne a subi une première chirurgie lors de la semaine 2 et une deuxième lors de la semaine 8. Malheureusement, cette personne ne portait pas son bracelet durant les jours qui ont suivi la première chirurgie et ne prenait plus part à la cueillette lors de la deuxième. C'est pourquoi l'analyse des changements comportementaux n'a pas pu être effectuée pour celle-ci. Les deux autres événements de vie se rapportent au participant numéro 2002 présenté plus tôt. Une gastro-entérite a été rapportée lors de la quatrième semaine (événement 1) et l'influenza lors de la huitième semaine (événement 2). L'analyse des changements comportementaux vis-à-vis ces événements a pu être effectuée partiellement, car la personne n'a pas porté le bracelet pendant la convalescence liée à ces deux événements. La Figure 5.34 reprend les informations sur les catégories d'actions effectuées par ce participant lors de la cueillette de données présentée précédemment et y ajoute des marqueurs aux endroits où les événements ont été rapportés.

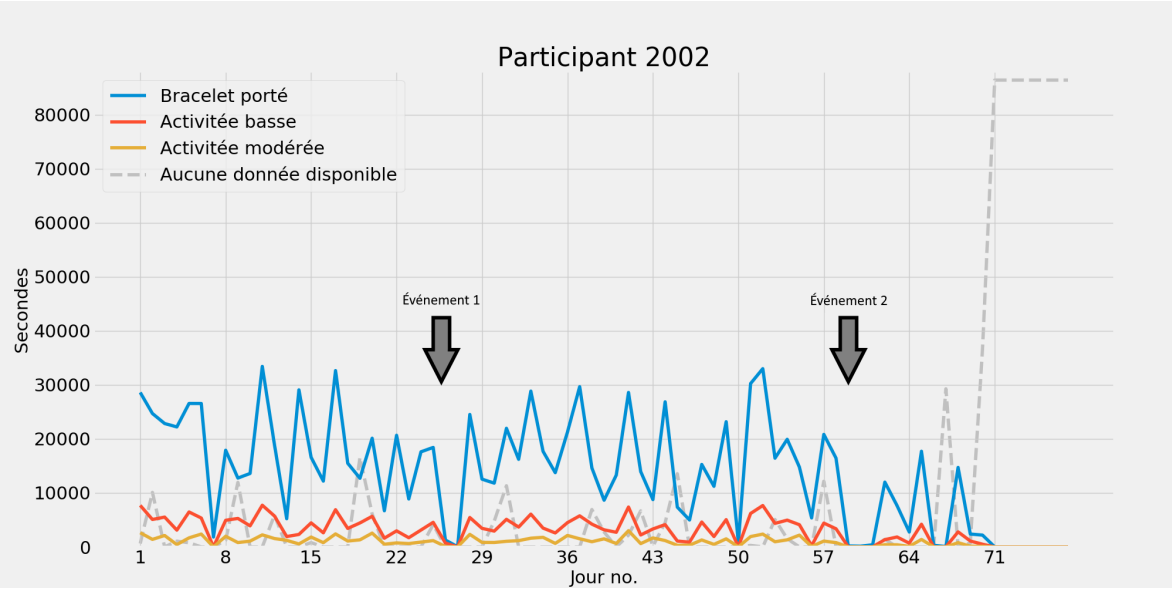


Figure 5.34 : Aperçu du temps écoulé pour chaque catégorie d'action effectuée par le participant numéro 2002 incluant les événements de vie

Afin d'évaluer l'impact de ces événements de vie sur la routine de l'individu, les résultats de l'étude des changements comportementaux présentée à la section 5.2.5 créée avec la semaine no. 1 comme semaine de base ont été utilisés. Certaines opérations supplémentaires sur ces données ont été effectuées afin d'évaluer si un changement de comportement plus important a été observé pour les semaines comportant un événement de vie.

— Premièrement, les scores de changement (*CS*) obtenus pour chaque caractéristique se rapportant à une semaine ont été combinés. Pour ce faire, la moyenne des scores de changement des trois caractéristiques a été calculée puis utilisée pour la semaine en question (score de changement combiné).

— Deuxièmement, la moyenne des scores de changement combinés de chaque semaine a

été calculée. Les scores de changement de la semaine 4 et de la semaine 9 ont ensuite été mis en relation avec celle-ci afin d'évaluer si le niveau de changement observé lors de ces semaines est supérieur à la moyenne.

Le Tableau 5.2 expose la moyenne des scores de changement combinés obtenue pour chaque méthode ainsi que la différence entre cette moyenne et le score de changement combiné obtenu pour la semaine 4 et la semaine 9. De plus, les Figures 5.35 et 5.36 démontrent ces données sous forme de graphique. Lorsque la différence avec la moyenne est négative, cela implique que le score de changement combiné est inférieur à la moyenne. Le contraire est aussi vrai : une différence positive implique un score plus élevé que la moyenne.

Tableau 5.2 : Analyse des différences entre la moyenne et les scores de changement combinés pour la semaine 4 et la semaine 9 du participant 2002

Méthode	**Moyenne des *CS* combinés**	**Différence semaine 4**	**Différence semaine 9**
sw-PCAR	0,82	-0,05	-0,28
RulSIF-vect	0,33	-0,12	0,73
RulSIF-mat	1,74	-1,13	0,01
RulSIF-mat-fea	1,74	-0,43	-0,17

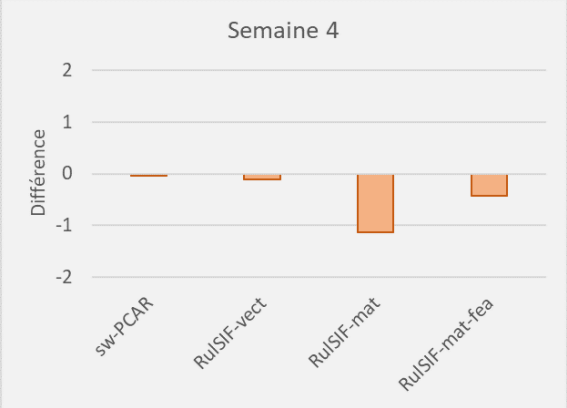


Figure 5.35 : Analyse des différences entre les scores de changement combinés et la moyenne pour la semaine 4 du participant 2002

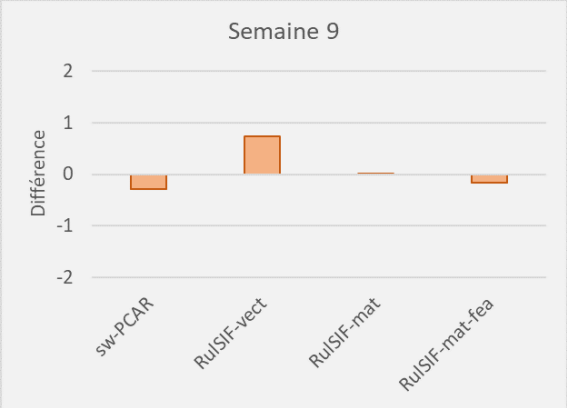


Figure 5.36 : Analyse des différences entre les scores de changement combinés et la moyenne pour la semaine 9 du participant 2002

Comme il est possible de la constater, les différences impliquent généralement un score moins élevé que la moyenne pour la semaine 4 et la semaine 9. Seulement la méthode RulSIF-vecteur appliquée à la neuvième semaine produit un changement supérieur à la moyenne. Or, c'est le seul indiquant une différence positive pour cette semaine et celle-ci est relativement basse. On peut donc conclure que les événements de vie n'ont pas affecté la routine de ces semaines de manière considérable.

Il faut tout de même porter attention à quelques facteurs qui ont atténué significativement l'effet de ces événements de vie sur la routine perçue pour ces deux semaines :

— Comme mentionné plus tôt, le participant 2002 a très peu porté le bracelet lors des journées suivant ces événements de vie.

— La routine de cet individu se situe dans les trois routines les plus instables observées (voir Figure 5.33). Cette instabilité peut causer du bruit et ainsi atténuer l'impact des

événements de vie enregistrés.

Il aurait été intéressant d'appliquer l'analyse des événements de vie sur un participant dont les données ne comportent pas ces facteurs atténuants.

5.3.4 LIMITES DE L'ÉTUDE

L'étude présentée dans ce mémoire comporte certaines limites. Comme il a été discuté plus tôt, l'étude des changements comportementaux a pu être effectuée sur seulement six personnes pour cause d'abandon de certains participants en cours de route. De plus, certaines contraintes ont fixé la durée de la cueillette de données sur dix semaines. Or, cette durée relativement courte limite le choix des algorithmes pouvant être appliqués sur l'analyse des données. Par exemple, il n'a pas été possible d'appliquer l'algorithme PCAR qui aurait permis de construire et de modéliser les courbes d'activités quotidiennes.

Bien que la technologie d'assistance à domicile permet la supervision du programme d'entraînement prescrit et assiste les participants dans la réalisation de leurs entraînements à domicile (Chapron *et al.*, 2021b), aucun contrôle sur la cueillette quotidienne de données n'est effectué. En effet, les éléments de la cueillette comme le port du bracelet ou le chargement de celui-ci sont laissés à la discrétion de chaque participant. Bien entendu, de la sensibilisation a été effectuée auprès de ces participants, mais nous ne pouvons ni surveiller ni forcer les participants à participer à la cueillette de manière assidue. D'ailleurs, comme il a été démontré dans la Section 5.3.1 de ce mémoire, la participation a été relativement faible pour certains participants.

D'autres enjeux ont dû être pris en compte lors de cette étude. La technologie utilisée pour la cueillette de données a été développée dans nos laboratoires spécialement pour ce projet. Cependant, elle comporte certaines limites. De plus, certains enjeux liés à la clientèle visée par la cueillette ont aussi dû être pris en compte.

Enjeux technologiques

Bien que la technologie de cueillette de données a été développée spécifiquement pour ce projet, certains compromis ont dû être effectués lors du développement. Le bracelet à centrale inertielle devait fonctionner avec une batterie et être portable. Pour répondre à ces contraintes, la technologie de communication sans-fil BLE a été utilisée de pair avec une plateforme de développement conçue spécifiquement pour être petite tout en fournissant cette connectivité (BLE Nano 2). Ces choix technologiques posent certains enjeux dont l'effet a été ressenti lors de la cueillette de données :

— Comme il a été souligné dans la Section 4.2.2 de ce mémoire, le microcontrôleur BLE Nano 2 possède une quantité limitée de mémoire interne. Cette limite empêche la sauvegarde des informations temporelles des activités effectuées lorsque le bracelet est en mode hors-ligne et diminue la qualité de l'information obtenue lorsque le bracelet n'est plus connecté à l'unité centrale de traitement. Pour pallier à cette limitation, les développements futurs pourraient inclure l'utilisation d'un stockage externe telle une carte mémoire (ex. : de type Micro SD). Une autre solution pourrait aussi impliquer

l'utilisation d'un lien entre le téléphone intelligent du participant et le bracelet afin de sauvegarder l'information lorsque celui-ci n'est plus à portée de la centrale.

— La technologie BLE est conçue pour des communications à courte distance. C'est d'ailleurs une limitation qui a été observée lors de ce projet. En effet, les données de certains participants qui demeurent dans une grande maison à plus d'un étage comportent un grand nombre de pertes de connexion avec l'unité centrale de traitement (ex. : participant 907). Ces pertes de connexions fréquentes impliquent une perte de données légère et cause du bruit dans les données. Des travaux futurs pourraient avoir comme objectif de modérer ces limitations en intégrant des dispositifs relais aux endroits stratégiques d'un domicile trop grand (BLE 5.0, Bluetooth Mesh) ou bien en étudiant la possibilité d'utiliser d'autres technologies de communication sans-fil conçues pour de plus longues distances telles LoRA, WiFi, etc.

Finalement, les dispositifs utilisés lors de la cueillette de données, bien que testés raisonnablement, restent des prototypes développés en laboratoire. De ce fait, ils sont généralement moins stables que les dispositifs commerciaux reconnus. Des problèmes techniques peuvent apparaître lors de la cueillette de données telles des bris matériels. Par exemple, sur les six participants dont l'analyse des changements comportementaux a été effectuée, deux bris matériels ont été signalés. Une personne a subi un bris au niveau de l'unité centrale de traitement et une autre personne a subi un bris au niveau du bracelet. Ces deux événements ont causé une perte de données plus ou moins importante selon le cas, ce qui a complexifié l'application des algorithmes de détection de changement comportemental.

Clientèle visée

La clientèle visée par ce projet est atteinte de DM1. La cueillette de données et l'analyse des changements comportementaux chez les individus atteint de dystrophie, ou même souffrant d'une maladie en général, comportent certaines difficultés supplémentaires.

Premièrement, les individus souffrant de DM1 sont généralement moins actifs en raison de leurs problèmes musculaires (Johnson, 2019). Ce taux d'activité plus bas peut atténuer l'effet de certains changements dans leur routine générale (ex. : participant 2002, effet de l'événement de vie influenza).

Deuxièmement, comme discuté lors de l'élaboration de l'hypothèse H3 (voir Section 1.2.2), les personnes atteintes de DM1 sont plus à risques de subir un événement de santé qui pourrait influencer le déroulement de la cueillette de données. Par exemple, un participant a subi deux opérations lors de la cueillette, ce qui a causé son retrait et invalidé ses données.

Étude comparative

La comparaison des différents algorithmes utilisés aurait été intéressante. En effet, évaluer la sensibilité de chacun vis-à-vis des événements rapportés aurait permis de comparer la performance des différents algorithmes dans le contexte de cette étude. Or, au vu des différents enjeux présentés plus tôt (durée de la cueillette, nombre de participants, assiduité, etc.), cette étude comparative n'était pas possible.

CONCLUSION

Cette partie expose les conclusions de ce mémoire. Premièrement, le travail présenté dans ce mémoire sera résumé. Deuxièmement, une courte revue des contributions sera faite. Troisièmement, les travaux futurs envisagés seront mentionnés. Finalement, cette section se terminera avec une conclusion personnelle.

Le sujet présenté porte sur l'analyse des changements comportementaux de personnes atteintes de DM1 effectuant un programme d'entraînement à domicile. Le chapitre 1 a permis de situer ces travaux de recherches et de présenter les différents termes nécessaires à la compréhension de ce travail de recherche. Le chapitre 2 consiste en la revue de littérature où il a été possible d'en savoir plus sur les principales méthodes du domaine de recherche et sur les différents contextes dans lesquels la détection de changement comportemental a été expérimentée. Le chapitre 3 a présenté la technologie d'assistance à domicile qui a été utilisée pour la cueillette des données ainsi que le déroulement de cette cueillette en situation réelle. Le chapitre 4 a ensuite décrit les différentes opérations qui ont été nécessaires pour transformer les données afin de permettre l'extraction de l'information utile pour la détection de changement comportemental. Cela a permis l'exécution des différents algorithmes de détection de changements qui ont été expliqués dans le Chapitre 5. Ensuite, il a été possible d'apprécier les résultats et d'effectuer certaines études de cas intéressante. Ce projet était assez ambitieux par le fait qu'il comportait de nombreux défis à surmonter. Malgré cela, il a été possible de recueillir des données et de faire l'analyse des changements comportementaux de

plusieurs participants. Aussi, comme plusieurs enjeux sont apparus tout au long du projet, une section sur les enjeux devant être pris en considération lors de telles expérimentations a été incluse dans le Chapitre 5.

REVUE DES CONTRIBUTIONS

Les contributions apportées par les travaux présentés dans ce mémoire sont multiples.

Tout d'abord, mentionnons que c'est la première fois qu'une solution d'assistance pour les entraînements à domicile de personnes atteintes de DM1 est expérimentée en contexte réel tout en recueillant des données qui ont permis l'analyse de l'évolution de l'activité physique. Cela a permis une contribution sur plus d'un volet. Premièrement, les données recueillies pourront être mises à la disposition de d'autres chercheurs. Deuxièmement, les méthodes sw-PCAR et RulSIF ont été utilisées dans un nouveau contexte, ce qui a permis d'en apprécier les particularités. Finalement, la cueillette de données dans ce contexte comporte plusieurs enjeux qui ont été expliqués.

Une autre contribution importante provient de la création du dispositif PIR+. Ce dispositif est un ajout non négligeable à la solution technologique à disposition du LIARA et pourra être utilisé lors de nouvelles cueillettes de données. Par exemple, ce dispositif pourra être mis à contribution dans de nouveaux déploiements liées au projet Acti-DM1 ou même dans d'autres projets.

TRAVAUX FUTURS

La Section 5.3.4 a fait part des limites de l'étude et des différents enjeux technologiques rencontrées. Il a été question de la portabilité du bracelet à centrale inertielle et de sa conception en laboratoire. Bien que la production du dispositif au sein du LIARA comporte certains avantages (ex. : plus grande flexibilité), l'utilisation d'un bracelet à centrale inertielle intégrée tels les bracelets ou les montres commerciales serait un atout pour la portabilité, la durabilité et la durée de charge de la batterie. C'est pourquoi l'utilisation d'un tel dispositif afin de remplacer le bracelet à centrale inertielle développé en laboratoire pourrait faire l'objet de travaux futurs.

Aussi, pour répondre aux enjeux de distance de communication de la technologie BLE tel que discuté plus tôt, l'utilisation de la technologie Bluetooth Mesh [16] pourrait être envisagée. En effet, l'implémention de cette fonctionnalité au niveau des dispositifs BLE installés dans le domicile à des endroits stratégiques (ex. : bracelet, PIR+), pourrait augmenter considérablement la couverture Bluetooth à l'intérieur de la résidence des patients.

Finalement, tel que discuté dans la Section 5.3.4, l'association du bracelet à un téléphone intelligent pourrait être envisagée afin d'adresser l'enjeu des déconnexions et de la sauvegarde des données temporelles lors d'une sortie du domicile.

16. https ://www.bluetooth.com/learn-about-bluetooth/recent-enhancements/mesh/

APPORT PERSONNEL

Le projet de recherche présenté dans ce mémoire comportait de nombreux défis que j'ai su relever au cours de ces deux dernières années. Mentionnons que je n'avais pratiquement aucune expérience en recherche avant d'entreprendre ce projet. C'est pourquoi j'ai acquis beaucoup de connaissances sur l'analyse des changements, sur la rédaction et sur la méthodologie de la recherche.

Cette expérience confirme encore une fois mon intérêt en recherche et je compte bien la mettre à contribution en continuant sur cette voie.

BIBLIOGRAPHIE

Adams, R. P. & MacKay, D. J. C. (2007). Bayesian Online Changepoint Detection. *arXiv preprint arXiv :0710.3742*. Récupéré à https://arxiv.org/abs/0710.3742

Aldehag, A., Jonsson, H., Lindblad, J., Kottorp, A., Ansved, T. & Kierkegaard, M. (2013). Effects of Hand-Training in Persons with Myotonic Dystrophy Type 1–a Randomised Controlled Cross-over Pilot Study. *Disability and Rehabilitation*, *35*(21), 1798–1807. doi:10.3109/09638288.2012.754952

Aminikhanghahi, S. & Cook, D. J. (2017). A Survey of Methods for Time Series Change Point Detection. *Knowledge and Information Systems*, *51*(2), 339–367. doi:10.1007/s10115-016-0987-z

Aminikhanghahi, S., Wang, T. & Cook, D. J. (2019). Real-Time Change Point Detection with Application to Smart Home Time Series Data. *IEEE Transactions on Knowledge and Data Engineering*, *31*(5), 1010–1023. doi:10.1109/TKDE.2018.2850347

APIMS (2020). Portail d'information sur la maladie de Steinert (DM1). Repéré le 10 novembre 2020 à https://www.maladiedesteinert.info.

Bouchard, G., Roy, R., Declos, M., Mathieu, J. & Kouladjian, K. (1989). Origin and Diffusion of the Myotonic Dystrophy Gene in the Saguenay Region (Quebec). *The Canadian Journal of Neurological Sciences. Le journal canadien des sciences neurologiques*, *16(1)*, 119–122.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32.

Chapron, K., Bouchard, K. & Gaboury, S. (2021a). Real-Time Gait Speed Evaluation at Home in a Multi Residents Context. *Multimedia Tools and Applications*, *80*, 12932–12949. doi:10.1007/s11042-020-08962-y

Chapron, K., Lapointe, P., Lessard, I., Darmstadt-Belanger, H., Bouchard, K., Gagnon, C., Lavoie, M., Duchesne, E. & Gaboury, S. (2021b). Acti-DM1 : Monitoring the Activity Level of People with Myotonic Dystrophy Type 1 Through Activity and Exercise Recognition. *IEEE Access*, *9*, 49960–49973. doi:10.1109/ACCESS.2021.3068877

Chapron, K., Plantevin, V., Thullier, F., Bouchard, K., Duchesne, E. & Gaboury, S. (2018). A More Efficient Transportable and Scalable System for Real-Time Activities and Exercises Recognition. *Sensors (Switzerland)*, *18*(1). doi:10.3390/s18010268

Cheng, J. & Greiner, R. (2013). Comparing Bayesian network classifiers. *arXiv preprint arXiv :1301.6684*. Récupéré à https://arxiv.org/abs/1301.6684

Cleland, I., Han, M., Nugent, C., Lee, H., McClean, S., Zhang, S. & Lee, S. (2014). Evaluation of Prompted Annotation of Activity Data Recorded from a Smart Phone. *Sensors (Basel, Switzerland)*, *14*(9), 15861–15879. doi:10.3390/s140915861

Dawadi, P. N., Cook, D. J. & Schmitter-Edgecombe, M. (2016). Modeling Patterns of Activities using Activity Curves. *Pervasive and Mobile Computing*, *28*, 51–68. doi: 10.1016/j.pmcj.2015.09.007

Efron, B. & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. New York, NY : CRC press.

Feuz, K. D., Cook, D. J., Rosasco, C., Robertson, K. & Schmitter-Edgecombe, M. (2015). Automated Detection of Activity Transitions for Prompting. *IEEE Transactions on Human-Machine Systems*, *45*(5), 575–585. doi:10.1109/THMS.2014.2362529

Gagnon, C., Mathieu, J., Jean, S., Laberge, L., Perron, M., Veillette, S., Richer, L. & Noreau, L. (2008). Predictors of Disrupted Social Participation in Myotonic Dystrophy Type 1. *Archives of Physical Medicine and Rehabilitation*, *89*(7), 1246–1255. doi:10.1016/j.apmr.2007.10.049

Gagnon, C., Petitclerc, É., Kierkegaard, M., Mathieu, J., Duchesne, É. & Hébert, L. J. (2018). A 9-year Follow-Up Study of Quantitative Muscle Strength Changes in Myotonic Dystrophy Type 1. *Journal of Neurology*, *265*(7), 1698–1705. doi:10.1007/s00415-018-8898-4

Gao, Q.-B. & Wang, Z.-Z. (2007). Center-Based Nearest Neighbor Classifier. *Pattern Recognition*, *40*(1), 346–349.

Hido, S., Idé, T., Kashima, H., Kubo, H. & Matsuzawa, H. (2008). Unsupervised Change Analysis Using Supervised Learning. Dans T. Washio, E. Suzuki, K. M. Ting, & A. Inokuchi

(dir.). *Advances in Knowledge Discovery and Data Mining*, 148–159., Berlin, Heidelberg. Springer Berlin Heidelberg.

Johnson, N. E. (2019). Myotonic Muscular Dystrophies. *Continuum (Minneapolis, Minn.)*, *25*(6), 1682–1695. doi:10.1212/CON.0000000000000793

Kanamori, T., Hido, S. & Sugiyama, M. (2009). A Least-Squares Approach to Direct Importance Estimation. *The Journal of Machine Learning Research*, *10*, 1391–1445.

Kawahara, Y., Yairi, T. & Machida, K. (2007). Change-Point Detection in Time-Series Data Based on Subspace Identification. Dans *Seventh IEEE International Conference on Data Mining (ICDM 2007)*, 559–564. IEEE.

Kim, J. K., Song, H. S., Kim, T. S. & Kim, H. K. (2005). Detecting the Change of Customer Behavior Based on Decision Tree Analysis. *Expert Systems*, *22*(4), 193–205. doi:10.1111/j.1468-0394.2005.00310.x

Kullback, S. & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86.

Lapointe, P., Chapron, K., Bouchard, K. & Gaboury, S. (2020). A New Device to Track and Identify people in a Multi-Residents Context. *Procedia Computer Science*, *170*, 403–410. doi:10.1016/j.procs.2020.03.082

Lessard, I., Gaboury, S., Gagnon, C., Bouchard, K., Chapron, K., Lavoie, M., Lapointe, P. & Duchesne, E. (2021). Effects and Acceptability of an Individualized Home-Based 10-Week Training Program in Adults with Myotonic Dystrophy Type 1. *Journal of Neuromuscular Diseases*, *8*(1), 137–149. doi:10.3233/JND-200570

Maimon, O. & Rokach, L. (2005). Data Mining and Knowledge Discovery Handbook.

Malladi, R., Kalamangalam, G. P. & Aazhang, B. (2013). Online Bayesian Change Point Detection Algorithms for Segmentation of Epileptic Activity. *2013 Asilomar Conference on Signals, Systems and Computers*, 1833–1837.

Orngreen, M. C., Olsen, D. B. & Vissing, J. (2005). Aerobic Training in Patients with Myotonic Dystrophy Type 1. *Annals of Neurology*, *57*(5), 754–757. doi:10.1002/ana.20460

Ortiz Laguna, J., Olaya, A. G. & Borrajo, D. (2011). A Dynamic Sliding Window Approach for Activity Recognition. Dans J. A. Konstan, R. Conejo, J. L. Marzo, & N. Oliver (dir.). *User Modeling, Adaption and Personalization*, 219–230., Berlin, Heidelberg. Springer Berlin Heidelberg.

Pearson, K. (1900). On the Criterion That a Given System of Deviations From the Probable In the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen From Random Sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157–175. doi:10.1080/14786440009463897

Podsiadlo, D. & Richardson, S. (1991). The Timed "Up & Go" : A Test of Basic Functional Mobility for Frail Elderly Persons. *Journal of the American Geriatrics Society*, *39*(2), 142–148. doi:10.1111/j.1532-5415.1991.tb01616.x

Rabiner, L. & Juang, B. (1986). An Introduction to Hidden Markov Models. *IEEE ASSP Magazine*, *3*(1), 4–16.

Rakthanmanon, T., Keogh, E. J., Lonardi, S. & Evans, S. (2011). Time Series Epenthesis : Clustering Time Series Streams Requires Ignoring Some Data. Dans *2011 IEEE 11th International Conference on Data Mining*, 547–556. IEEE.

Randolph, C. (1998). *Repeatable Battery for the Assessment of Neuropsychological Status (RBANS).* San Antonio, TX : Psychological Corporation.

Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M. & Srivastava, M. (2010). Using Mobile Phones to Determine Transportation Modes. *ACM Trans. Sen. Netw.*, *6*(2). doi: 10.1145/1689239.1689243

Reeves, J., Chen, J., Wang, X. L., Lund, R. & Lu, Q. Q. (2007). A Review and Comparison of Changepoint Detection Techniques for Climate Data. *Journal of Applied Meteorology and Climatology*, *46*(6), 900–915. doi:10.1175/JAM2493.1

Salzberg, S. L. (1994). C4. 5 : Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, inc., 1993.

Shumway, R. H. & Stoffer, D. S. (2017). *Time Series Analysis and its Applications : with R Examples*. Springer.

Sprint, G., Cook, D., Weeks, D., Dahmen, J. & La Fleur, A. (2017). Analyzing Sensor-Based Time Series Data to Track Changes in Physical Activity during Inpatient Rehabilitation. *Sensors (Basel, Switzerland)*, *17*(10). doi:10.3390/s17102219

Sprint, G., Cook, D. J. & Fritz, R. (2021). Behavioral Differences Between Subject Groups Identified Using Smart Homes and Change Point Detection. *IEEE Journal of Biomedical and Health Informatics*, *25*(2), 559–567. doi:10.1109/JBHI.2020.2999607

Sprint, G., Cook, D. J., Fritz, R. & Schmitter-Edgecombe, M. (2016a). Using Smart Homes to Detect and Analyze Health Events. *Computer*, *49*(11), 29–37. doi:10.1109/MC.2016.338

Sprint, G., Cook, D. J. & Schmitter-Edgecombe, M. (2016b). Unsupervised Detection and Analysis of Changes in Everyday Physical Activity Data. *Journal of Biomedical Informatics*, *63*, 54–65. doi:10.1016/j.jbi.2016.07.020

Staudacher, M., Telser, S., Amann, A., Hinterhuber, H. & Ritsch-Marte, M. (2005). A New Method for Change-point Detection Developed for On-line Analysis of the Heart Beat Variability During Sleep. *Physica A : Statistical Mechanics and its Applications*, *349*(3), 582–596. doi:10.1016/j.physa.2004.10.026

Steinwart, I. & Christmann, A. (2008). *Support Vector Machines*. Springer Science & Business Media.

Sugiyama, M., Suzuki, T., Nakajima, S., Kashima, H., von Bünau, P. & Kawanabe, M. (2008). Direct Importance Estimation for Covariate Shift Adaptation. *Annals of the Institute of Statistical Mathematics*, *60*(4), 699–746. doi:10.1007/s10463-008-0197-x

Tollback, A., Eriksson, S., Wredenberg, A., Jenner, G., Vargas, R., Borg, K. & Ansved, T. (1999). Effects of High Resistance Training in Patients with Myotonic Dystrophy. *Scandinavian Journal of Rehabilitation Medicine*, *31*(1), 9–16.

Wei, L. & Keogh, E. (2006). Semi-Supervised Time Series Classification. Dans *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '06, 748–753., New York, NY, USA. Association for Computing Machinery. doi:10.1145/1150402.1150498

Yamada, M., Suzuki, T., Kanamori, T., Hachiya, H. & Sugiyama, M. (2013). Relative Density-Ratio Estimation for Robust Distribution Comparison. *Neural Computation*, *25*, 1324–1370. doi:10.1162/NECO_a_00442

Zakaria, J., Mueen, A. & Keogh, E. (2012). Clustering Time Series Using Unsupervised-Shapelets. Dans *2012 IEEE 12th International Conference on Data Mining*, 785–794. IEEE.

APPENDICE A

CERTIFICATS D'APPROBATION ÉTHIQUE

A.1 APPROBATION ÉTHIQUE 2019-035

Centre intégré
universitaire de santé
et de services sociaux
du Saguenay–
Lac-Saint-Jean
Québec

Direction de la recherche

Le 25 juin 2019

Madame Elise Duchesne, Ph. D.
Université du Québec à Chicoutimi
555, boulevard de l'Université, H2-1030
Chicoutimi (Québec) G7H 2B1

Objet : Approbation éthique finale du Comité d'éthique de la recherche du CIUSSS du Saguenay–Lac-Saint-Jean

Titre du projet : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle

Numéro de projet attribué par le CIUSSS du Saguenay–Lac-Saint-Jean : 2019-035

Madame,

Le Comité d'éthique de la recherche du Centre intégré universitaire de santé et de services sociaux du Saguenay-Lac-Saint-Jean a évalué votre projet de recherche à sa réunion tenue le 25 juin 2019. Lors de cette réunion, le comité a examiné les documents suivants :

1. Lettre d'approbation scientifique finale, datée du 7 juin 2019
2. Lettre de présentation, datée du 23 mai 2019
3. Formulaire de demande d'évaluation éthique, signé le 23 mai 2019
4. Protocole de recherche intitulé : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle, daté du 23 mai 2019
5. Formulaire d'information et de consentement, daté du 19 juin 2019
6. Questionnaire d'informations générales, daté du 3 mai 2019
7. Avis d'octroi du regroupement stratégique INTER, daté du 8 janvier 2019
8. Budget
9. Lettre d'appui de la chef en réadaptation déficience physique, Jonquière, datée du 16 mai 2019
10. Formulaire d'étude de la convenance institutionnelle, signé le 1er mai 2019
11. Document – Demande de consultation de dossiers pour fins d'étude, signé le 9 avril 2019
12. Curriculum vitae de :
    a. Benjamin Gallais, Ph. D.
    b. Kévin Bouchard, Ph. D.
    c. Sébastien Gaboury, Ph. D.
13. Document UQAC – Autorisation – Accès au dossier, signé le 3 mai 2019
14. Document UQAC – Autorisation – Dépôt d'une demande de certification éthique, signé par le doyen le 27 mai 2019

Affilié
UNIVERSITÉ DE MONTRÉAL
UNIVERSITÉ DE SHERBROOKE

305, rue Saint-Vallier, case postale 266
Chicoutimi (Québec) G7H 5H6
Téléphone : 418 541-1234, poste 2369
Télécopieur : 418 541-1124
www.santesaglac.com

Le Comité d'éthique de la recherche a évalué en comité plénier les aspects éthiques de votre projet de recherche.

Rappelons d'entrée de jeu que votre projet de recherche a fait l'objet d'une évaluation préliminaire. Lors de cette évaluation préliminaire, nous avons apporté des modifications au formulaire d'information et de consentement. Suite à cette évaluation préliminaire, vous nous avez fait parvenir le 19 juin 2019, la nouvelle version du formulaire d'information et de consentement pour la participation de majeurs.

Étant donné que les commentaires et les modifications avaient déjà été intégrés au document et que nous n'avions pas d'autres commentaires à formuler suite à notre réunion, nous avons donc le plaisir de vous informer que votre projet de recherche a été approuvé, tel quel, à l'unanimité par le Comité.

Les documents que le Comité d'éthique de la recherche du CIUSSS du Saguenay–Lac-Saint-Jean a approuvés et que vous pouvez utiliser pour la réalisation de votre projet sont les suivants :

1. Protocole de recherche intitulé : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle, daté du 25 juin 2019
2. Formulaire d'information et de consentement, daté du 25 juin 2019
3. Questionnaire d'informations générales, daté du 25 juin 2019

Cette approbation éthique est valide pour un an à compter du 25 juin 2019, date de l'approbation éthique finale. Cette décision peut être suspendue ou révoquée en cas de non-respect de ces exigences.

Un mois avant la date d'échéance vous devrez faire une demande de renouvèlement auprès du Comité d'éthique de la recherche du CIUSSS du Saguenay–Lac-Saint-Jean en utilisant le document du Comité prévu à cet effet.

Si un mois après l'échéance le Comité n'a rien reçu, il procèdera à la fermeture du projet de recherche sans autre avis. Ceci signifie que vous devrez recommencer tout le processus d'évaluation et d'autorisation pour le projet ciblé si vous souhaitez continuer le projet.

Dans le cadre du suivi continu, le Comité vous demande de vous conformer aux exigences suivantes en utilisant les formulaires du Comité prévus à cet effet :

1. De soumettre toute demande de modification au projet de recherche ou à tout document approuvé par le Comité pour la réalisation de votre projet.
2. De soumettre, dès que cela est porté à votre connaissance, tout nouveau renseignement ou toute modification à l'équilibre clinique susceptible d'affecter l'intégrité ou l'éthicité du projet de recherche, d'accroître les risques et les inconvénients pour les participants, de nuire au bon déroulement du projet ou d'avoir une incidence sur le désir d'un participant de continuer à participer au projet.
3. De soumettre, dès que cela est porté à votre connaissance et en lien avec la réalisation de ce projet, tout accident survenu dans votre site.
4. De soumettre, dès que cela est porté à votre connaissance, l'interruption prématurée du projet de recherche, qu'elle soit temporaire ou permanente.
5. De soumettre, dès que cela est porté à votre connaissance, tout problème constaté à la suite d'une activité de surveillance ou de vérification menée par un tiers et susceptible de remettre en question l'intégrité ou l'éthicité du projet de recherche

6. De soumettre, dès que cela est porté à votre connaissance, toute suspension ou annulation de l'approbation octroyée par un organisme de subvention ou de réglementation.
7. De soumettre, dès que cela est porté à votre connaissance, toute procédure en cours de traitement d'une plainte ou d'une allégation de manquement à l'intégrité ou à l'éthicité ainsi que des résultats de la procédure.
8. De soumettre, toute déviation au projet de recherche susceptible d'augmenter le niveau de risque ou susceptibles d'influer sur le bien-être du participant ou d'entacher le consentement du participant.
9. De soumettre une demande de renouvellement annuel de l'approbation du projet de recherche.
10. De soumettre le rapport de la fin du projet de recherche.

Nous vous informons que nous acheminerons l'approbation éthique finale de votre projet à la personne formellement mandatée responsable au CIUSSS du Saguenay–Lac-Saint-Jean pour que vous puissiez obtenir l'autorisation de réaliser votre projet dans les murs de l'établissement.

Nous vous avisons également que vous ne pouvez commencer votre projet avant d'avoir obtenu l'autorisation de la personne formellement mandatée responsable au CIUSSS du Saguenay–Lac-Saint-Jean.

De plus, nous vous rappelons que vous devez conserver pour une période d'au moins un an suivant la fin du projet, un répertoire distinct comprenant les noms, prénoms, coordonnées, date du début et de fin de la participation de chaque sujet de recherche.

Nous vous demandons d'utiliser pour votre projet le numéro de référence suivant 2019-035 pour toute correspondance avec le Comité d'éthique de la recherche.

Nous vous confirmons qu'aucune des personnes qui ont procédé à l'évaluation n'était placée dans une situation de conflit d'intérêts.

Le Comité d'éthique de la recherche du Centre intégré universitaire de santé et de services sociaux du Saguenay-Lac-Saint-Jean est désigné par le ministre de la Santé et des Services sociaux aux fins de l'application de l'article 21 du Code civil du Québec et suit les règles émises par l'Énoncé de politique des trois conseils et les Bonnes pratiques cliniques telles qu'élaborées par la Conférence internationale sur l'harmonisation (CIH).

Veuillez recevoir l'expression de nos sentiments les meilleurs.

Johane de Champlain, vice-présidente substitut
Comité d'éthique de la recherche
CIUSSS du Saguenay–Lac-Saint-Jean
JdeC/ah

c.c. Isabelle Côté, professionnelle de recherche : isabelle.cote7@usherbrooke.ca

Dr Martin Fortin, M.D., personne formellement mandatée responsable CIUSSS du Saguenay–Lac-Saint-Jean

A.2 APPROBATION ÉTHIQUE 2019-035 (RENOUVELLEMENT)

Centre intégré universitaire de santé et de services sociaux du Saguenay–Lac-Saint-Jean
Québec

Le 28 mai 2020

Madame Elise Duchesne, Ph. D.
Université du Québec à Chicoutimi
555, boulevard Université local H2-1030
Chicoutimi (Québec) G7H 2B1

Objet : Renouvellement annuel de l'approbation éthique de votre projet au CIUSSS du Saguenay-Lac-Saint-Jean - R1 - formulaire F9-756

Titre : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle

Numéro du projet : 2019-035 Entrainement à domicile

Identifiant Nagano : 2020-7

Madame,

Vous avez soumis au Comité d'éthique de la recherche du Centre intégré universitaire de santé et de services sociaux du Saguenay-Lac-Saint-Jean, le 27 mai 2020, une demande de renouvellement de l'approbation éthique à l'égard de votre projet.

Votre demande de renouvellement a fait l'objet d'une évaluation. Suite à cette évaluation, le tout étant jugé satisfaisant, j'ai le plaisir de vous informer que votre demande de renouvellement a été approuvée par le Comité d'éthique de la recherche du CIUSSS du Saguenay–Lac-Saint-Jean.

Ainsi, vous pouvez poursuivre votre étude pour un an, et ce, à compter du 25 juin 2020 jusqu'au 25 juin 2021.

UN mois avant la date d'échéance, vous devrez faire une demande de renouvellement auprès du Comité d'éthique de la recherche CIUSSS du Saguenay–Lac-Saint-Jean en utilisant le formulaire de demande de renouvellement annuel - formulaire 9 prévu à cet effet dans Nagano.

Si un mois après la date d'échéance le Comité n'a rien reçu, il procédera à la fermeture du projet de recherche sans autre avis. Dans l'éventualité où vous voudriez réaliser ce projet de recherche, vous devrez le soumettre à nouveau. Ce dernier fera l'objet d'une nouvelle évaluation réalisée par le Comité d'éthique de la recherche et le Comité de la convenance institutionnelle et il devra par la suite faire l'objet d'une autorisation émise par la personne formellement mandatée.

Nous vous rappelons que dans le cadre de son suivi continu, le Comité vous demande de vous conformer aux exigences suivantes en utilisant les formulaires du Comité prévus à cet effet :

- De soumettre toute demande de modification au projet de recherche ou à tout document approuvé par le Comité pour la réalisation de votre projet.
- De soumettre, dès que cela est porté à votre connaissance, tout nouveau renseignement ou toute modification à l'équilibre clinique susceptible d'affecter l'intégrité ou l'éthicité du projet de recherche, d'accroître les risques et les inconvénients pour les participants, de nuire au bon déroulement du projet ou d'avoir une incidence sur le désir d'un participant de continuer à participer au projet.
- De soumettre, dès que cela est porté à votre connaissance et en lien avec la réalisation de ce projet, tout accident survenu dans votre site.
- De soumettre, dès que cela est porté à votre connaissance, l'interruption prématurée du projet de recherche,

qu'elle soit temporaire ou permanente.

- De soumettre, dès que cela est porté à votre connaissance, tout problème constaté à la suite d'une activité de surveillance ou de vérification menée par un tiers et susceptible de remettre en question l'intégrité ou l'éthicité du projet de recherche
- De soumettre, dès que cela est porté à votre connaissance, toute suspension ou annulation de l'approbation octroyée par un organisme de subvention ou de réglementation.
- De soumettre, dès que cela est porté à votre connaissance, toute procédure en cours de traitement d'une plainte ou d'une allégation de manquement à l'intégrité ou à l'éthicité ainsi que des résultats de la procédure.
- De soumettre, toute déviation au projet de recherche susceptible d'augmenter le niveau de risque ou susceptibles d'influer sur le bien-être du participant ou d'entacher le consentement du participant.
- De soumettre une demande de renouvellement annuel de l'approbation du projet de recherche.
- De soumettre le rapport de la fin du projet de recherche.

De plus, nous vous rappelons que vous devez conserver pour une période d'au moins un an suivant la fin du projet, un répertoire distinct comprenant les noms, prénoms, coordonnées, date du début et de fin de la participation de chaque sujet de recherche.

Finalement, nous vous mentionnons que la présente décision vaut pour une année et peut être suspendue ou révoquée en cas de non-respect de ces exigences.

Le Comité d'éthique de la recherche du Centre intégré universitaire de santé et de services sociaux du Saguenay-Lac-Saint-Jean est désigné par le ministre de la Santé et des Services sociaux aux fins de l'application de l'article 21 du Code civil du Québec et suit les règles émises par l'Énoncé de politique des trois conseils et les Bonnes pratiques cliniques de la CIH.

Recevez l'expression de nos sentiments les meilleurs.

Madame Johane de Champlain

Vice-présidente

Comité d'éthique de la recherche

CIUSSS du Saguenay-Lac-Saint-Jean

A.3 APPROBATION ÉTHIQUE UQAC 2019-035

UQAC
Université du Québec
à Chicoutimi

Le 22 août 2019

À l'attention de :

Sébastien Gaboury, Professeur
Département d'informatique et de mathématique
Université du Québec à Chicoutimi

Élise Duchesne, Professeure
Département des sciences de la santé
Université du Québec à Chicoutimi

Titre du projert de recherche : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle

 # Projet : 2020-223 (CIUSSS - 2019-035)

Financement : Subvention du regroupement stratégique INTER – Ingénierie de technologies interactives en réadaptation, financé par le Fonds de recherche du Québec nature et technologies (FRQNT)

Objet : Approbation éthique de votre projet de recherche

Bonjour,

Votre projet de recherche a fait l'objet d'une reconnaissance en matière d'éthique de la recherche avec des êtres humains par le Comité d'éthique de la recherche de l'Université du Québec à Chicoutimi (CER-UQAC) conformément à l'Entente entre le CIUSSS-SLSJ et l'UQAC.

Informations concernant le projet appouvé par le CER du CIUSSS-SLSJ:

- Titre du projet : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle
- Numéro du projet attribué par le CER du CIUSSS-SLSJ: 2019-035
- Chercheur principal: Élise Duchesne

Un certificat d'approbation éthique qui atteste de la conformité de votre projet de recherche à laPolitique d'éthique de la recherche avec des êtres humains de l'UQAC est émis en date du 22 août 2019. Prenez note que ce certificat est valide jusqu'au **22 août 2020**.

Il est de la responsabilité des chercheurs d'assurer que leurs projets de recherche conservent une approbation éthique pour toute la durée des travaux de recherche et d'informer le CER de la fin de ceux-ci. Vous devrez donc obtenir le renouvellement de votre approbation éthique avant l'expiration de ce certificat à l'aide du formulaire *F7 - Renouvellement annuel*. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre certificat. Un rapport non remis entraînera la fermeture du dossier éthique dans les 30 jours qui suivent et une mention y sera ajoutée. Si votre projet est terminé, vous devrez remplir le formulaire *F9 - Fin de projet*.

Enfin, puisque votre demande indique l'obtention d'un financement, le Décanat de la recherche et de la création est mis en copie conforme afin de l'informer de l'obtention de la certification éthique.

Exporté le 2021-04-10 11:21 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-0b3ca769-69a9-44be-8d11-560c426a5748http://uqac.nagano.ca/verification/uqac-0b3ca769-69a9-44be-8d11-560c426a5748

UQAC
Université du Québec
à Chicoutimi

En vous souhaitant bon succès dans la réalisation de votre recherche, veuillez recevoir nos salutations distinguées.

Le CER-UQAC

Exporté le 2021-04-10 11:21 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-0b3ca769-69a9-44be-8d11-560c426a5748http://uqac.nagano.ca/verification/uqac-0b3ca769-69a9-44be-8d11-560c426a5748

UQAC
Université du Québec
à Chicoutimi

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une reconnaissance en matière d'éthique de la recherche avec des êtres humains selon l'Entente entre le CIUSSS-SLSJ et l'UQAC.

Projet # : 2020-223 (CIUSSS - 2019-035)

Titre du projet de recherche : Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle

Chercheur principal :
Sébastien Gaboury, Professeur
Département d'informatique et de mathématique, Université du Québec à Chicoutimi

Cochercheur :
Élise Duchesne, Professeure
Département des sciences de la santé, Université du Québec à Chicoutimi

Financement : Subvention du regroupement stratégique INTER – Ingénierie de technologies interactives en réadaptation, financé par le Fonds de recherche du Québec nature et technologies (FRQNT)

Date d'approbation du projet : 22 août 2019

Date d'entrée en vigueur du certificat : 22 août 2019

Date d'échéance du certificat : 22 août 2020

Stéphane Allaire

Exporté le 2021-04-10 11:21 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-0b3ca769-69a9-44be-8d11-560c426a5748http://uqac.nagano.ca/verification/uqac-0b3ca769-69a9-44be-8d11-560c426a5748

A.4 APPROBATION ÉTHIQUE UQAC 2019-035 (RENOUVELLEMENT)

UQAC
Université du Québec
à Chicoutimi

Le 22 juin 2020

RENOUVELLEMENT DE L'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'un renouvellement de l'approbation éthique émise par le CER-UQAC et qu'il satisfait aux exigences de la politique de l'Université du Québec à Chicoutimi en matière d'éthique de la recherche avec des êtres humains.

*** ***IMPORTANT-COVID-19:*** Tout projet impliquant des interactions en personnes **ne peut se poursuivre d'ici le 1er septembre 2020**, à moins d'avoir obtenu une autorisation spécifique du CER à cet effet. Cette consigne sera révisée au cours du mois d'août selon l'évolution des mesures de protection sanitaire émises par la Santé publique.***

Projet # : 2020-223, 2019-035 (CIUSSS-SLSJ), 2020-7 (CIUSSS-SLSJ)

Titre du projet de recherche: Développement d'une technologie d'assistance à domicile pour les personnes atteintes de dystrophie myotonique de type 1 visant à améliorer la capacité musculaire et fonctionnelle

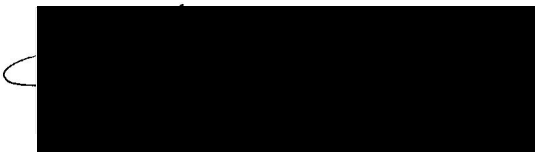
Chercheur principal à l'UQAC
Sebastien Gaboury, Professeur
département d'informatique et de mathématique, Université du Québec à Chicoutimi

Financement : FRQNT

Date de l'approbation éthique initiale du projet : 22 août 2019

Date du prochain renouvellement : 22 août 2021.
N.B. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre certificat afin de remplir le formulaire F7 - Renouvellement annuel.

- Si votre projet se termine avant la date du prochain renouvellement, vous devrez remplir le formulaire ***F9 - Fin de projet***.
- Si des modifications sont apportées à votre projet avant l'échéance du certificat, vous devrez remplir le formulaire ***F8 - Modification de projet***.
- Tout nouveau membre de votre équipe de recherche devra être déclaré au CER-UQAC lors de votre prochaine demande de renouvellement ou lors de la fin de votre projet si le renouvellement n'est pas requis. ATTENTION: Vous devez faire signer une déclaration d'honneur aux personnes ayant accès aux participants (ou à des données nominatives sur les participants) et la conserver dans vos dossiers de recherche.
- Si vous avez des cochercheurs dans d'autres universités, veuillez leur transmettre ce certificat.
- Si votre projet est financé, le Décanat de la recherche et de la création sera mis en copie conforme afin de l'informer du renouvellement de votre certification éthique.



Exporté le 2021-04-10 11:22 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-8d5008de-b241-4c9c-b249-46bee1e52822http://uqac.nagano.ca/verification/uqac-8d5008de-b241-4c9c-b249-46bee1e52822

UQAC
Université du Québec
à Chicoutimi

Stéphane Allaire

Exporté le 2021-04-10 11:22 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-8d5008de-b241-4c9c-b249-46bee1e52822http://uqac.nagano.ca/verification/uqac-8d5008de-b241-4c9c-b249-46bee1e52822

A.5 APPROBATION ÉTHIQUE 2019-220

UQAC
Université du Québec
à Chicoutimi

Le 21 mai 2019

À l'attention de :

Camille Bertuglia, Étudiant
Maîtrise en informatique
Département d'informatique et de mathématique

Kevin Bouchard, Professeur
Département d'informatique et de mathématique
Sebastien Gaboury, Professeur
Département d'informatique et de mathématique
Julien Maitre, Chargé de cours
Département d'informatique et de mathématique

Objet : Approbation éthique de votre projet de recherche

 # Projet : 2019-220

Titre du projet de recherche : Reconnaissance d'activités dans un habitat intelligent

Source de financement : Programme de soutien à la valorisation et au transfert (PSVT) du ministère de l'Économie et de l'Innovation (MESI) (financement de Kevin Bouchard)

Bonjour,

Votre projet de recherche a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains par le Comité d'éthique de la recherche de l'Université du Québec à Chicoutimi (CER-UQAC).

Un certificat d'approbation éthique qui atteste de la conformité de votre projet de recherche à la Politique d'éthique de la recherche avec des êtres humains de l'UQAC est émis en date du 21 mai 2019. Prenez note que ce certificat est **valide jusqu'au 21 mai 2020.**

Veuillez noter que le Formulaire d'information et de consentement a été revu et légèrement modifié par le CER-UQAC. Ce document sera déposé dans votre projet (voir les documents précédés d'un carré mauve dans la section "Fichiers" - "Documents officiels"). Cette version est celle autorisée par le CER et elle devra être utilisée pour votre projet.

Notez qu'en vertu de la Politique d'éthique de la recherche avec des êtres humains, il est de la responsabilité des chercheurs d'assurer que leurs projets de recherche conservent une approbation éthique pour toute la durée des travaux de recherche et d'informer le CER de la fin de ceux-ci. Vous devrez donc obtenir le renouvellement de votre approbation éthique avant l'expiration de ce certificat à l'aide du formulaire *F7 - Renouvellement annuel*. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre certificat.

Un rapport non remis entraînera la fermeture du dossier éthique dans les 30 jours qui suivent et une mention y sera ajoutée. Ainsi, la poursuite de la **cueillette de données** auprès des participants, sans certification éthique valide, ou le fait d'**apporter une modification significative** (à la population ciblée, au formulaire de consentement, au protocole d'expérimentation, à la méthode de collecte ou de traitement des données, etc.) **ou affectant le niveau de risque du projet** sans approbation du CER-UQAC représentent des situations relevant

Exporté le 2021-06-03 08:23 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-d2e73356-f015-49d9-b09e-f27560f1af6chttps://nagano.uqac.ca/verification/uqac-d2e73356-f015-49d9-b09e-f27560f1af6c

de la Politique relative à la conduite responsable en recherche et en création. De plus, le chercheur a l'obligation de signaler tout incident grave dès qu'il survient.

Si des modifications sont apportées à votre projet avant l'échéance du certificat, vous devrez remplir le formulaire *F8 - Modification de projet* et obtenir l'approbation du CER avant de mettre en oeuvre ces modifications. Si votre projet est terminé avant l'échéance du certificat, vous devrez remplir le formulaire *F9 - Fin de projet* .

Enfin, veuillez prendre note que le Décanat des études est mis en copie conforme de ce courriel afin de l'aviser de l'obtention de votre certification éthique. De plus, puisque votre demande d'approbation indique l'obtention d'un financement, le Décanat de la recherche et de la création est mis en copie conforme afin de l'informer de l'obtention de votre certification éthique.

En vous souhaitant bon succès dans la réalisation de votre recherche, veuillez recevoir mes salutations distinguées.

Tommy Chevrette,
Professeur et président du CER

Exporté le 2021-06-03 08:23 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-d2e73356-f015-49d9-b09e-f27560f1af6chttps://nagano.uqac.ca/verification/uqac-d2e73356-f015-49d9-b09e-f27560f1af6c

UQAC
Université du Québec
à Chicoutimi

CERTIFICAT D'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'une évaluation en matière d'éthique de la recherche avec des êtres humains et qu'il satisfait aux exigences de la politique de l'UQAC en cette matière.

Les membres jugent que ce projet rencontre les critères d'une recherche à risque minimal.

Projet # : 2019-220

Titre du projet de recherche : Reconnaissance d'activités dans un habitat intelligent

Chercheur principal :
Camille Bertuglia, Étudiant
Maîtrise en informatique
Département d'informatique et de mathématique

Directeur/codirecteur :
En provenance de l'UQAC:
Kevin Bouchard, Professeur
Département d'informatique et de mathématique
Sebastien Gaboury, Professeur
Département d'informatique et de mathématique
Julien Maitre, Chargé de cours
Département d'informatique et de mathématique

Financement : Programme de soutien à la valorisation et au transfert (PSVT) du ministère de l'Économie et de l'Innovation (MESI) (financement de Kevin Bouchard)

Date d'approbation du projet : 21 mai 2019

Date d'entrée en vigueur du certificat : 21 mai 2019

Date d'échéance du certificat : 21 mai 2020

[redacted]

Tommy Chevrette

Exporté le 2021-06-03 08:23 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-d2e73356-f015-49d9-b09e-f27560f1af6chttps://nagano.uqac.ca/verification/uqac-d2e73356-f015-49d9-b09e-f27560f1af6c

A.6 APPROBATION ÉTHIQUE 2019-220 (RENOUVELLEMENT)

UQAC
Université du Québec
à Chicoutimi

Le 10 mars 2020

RENOUVELLEMENT DE L'APPROBATION ÉTHIQUE

La présente atteste que le projet de recherche décrit ci-dessous a fait l'objet d'un renouvellement de l'approbation éthique émise par le CER-UQAC et qu'il satisfait aux exigences de la politique de l'Université du Québec à Chicoutimi en matière d'éthique de la recherche avec des êtres humains.

Projet # : 2019-220

Titre du projet de recherche
Reconnaissance d'activités dans un habitat intelligent

Chercheur principal à l'UQAC
Camille Bertuglia, étudiante
département d'informatique et de mathématique, UQAC

Direction / Codirection de recherche
En provenance de l'UQAC: Kevin Bouchard; Sebastien Gaboury; Julien Maitre

Cochercheur(s)
En provenance de l'UQAC: Kevin Chapron; Patrick Lapointe

Financement PSVT

Date de l'approbation éthique initiale du projet : 21 mai 2019

Date du prochain renouvellement : 21 mai 2021.
N.B. Un rappel automatique vous sera envoyé par courriel quelques semaines avant l'échéance de votre certificat afin de remplir le formulaire F7 - Renouvellement annuel.

- Si votre projet termine avant la date du prochain renouvellement, vous devrez remplir le formulaire *F9 - Fin de projet.*
- Si des modifications sont apportées à votre projet avant l'échéance du certificat, vous devrez remplir le formulaire *F8 - Modification de projet.*
- Tout nouveau membre de votre équipe de recherche devra être déclaré au CER-UQAC lors de votre prochaine demande de renouvellement ou lors de la fin de votre projet si le renouvellement n'est pas requis. ATTENTION: Vous devez faire signer une déclaration d'honneur aux personnes ayant accès aux participants (où à des données nomnatives sur les participants) et la conserver dans vos dossiers de recherche.
- Si vous avez des cochercheurs dans d'autres universités, veuillez leur transmettre ce certificat.
- Si votre projet est financé, le Décanat de la recherche et de la création sera mis en copie conforme afin de l'informer du renouvellement de votre certification éthique.

[redacted]

Exporté le 2021-06-03 08:24 par Gaboury, Sebastien --- CODE DE VALIDATION NAGANO: uqac-4410d31d-f5b7-4400-aa7c-98de1f62e0a4https://nagano.uqac.ca/verification/uqac-4410d31d-f5b7-4400-aa7c-98de1f62e0a4