

TESIS DOCTORAL

2015

**MÉTODO AUTOMÁTICO Y ADAPTATIVO PARA LA
DETECCIÓN DE ANOMALÍAS EN LA ACTIVIDAD
FÍSICA MEDIANTE UN SENSOR DE ACELERACIÓN
NO INVASIVO**

JUAN LUIS CARÚS CANDÁS

INGENIERO EN TELECOMUNICACIÓN

**PROGRAMA DE DOCTORADO EN TECNOLOGÍAS
INDUSTRIALES**

GABRIEL DÍAZ ORUETA

EDUARDO ÁLVAREZ ÁLVAREZ

A mis abuelos Manolo y Encarna,
por enseñarme a soñar

AGRADECIMIENTOS

A Víctor Manuel Peláez, porque su nombre también debería estar escrito en la portada. Sin su ingenio y su confianza, jamás hubiera sido capaz de abordar este trabajo.

A Gabriel Díaz y Eduardo Álvarez, directores de esta tesis, por su apoyo y paciencia. Gracias por darme la mano en este camino y ayudarme a mantener el equilibrio.

A todo el equipo de Fundación CTIC – Centro Tecnológico por confiar en mí y permitir mi desarrollo profesional en un clima de trabajo excepcional. Gracias a todas y cada una de las personas que forman y formaron parte del centro, en especial al grupo de investigación 4U.

A mi familia por apoyar sin preguntar, por nunca pedir explicaciones y por hacerme sentir siempre capaz de todo aquello que me proponga. En especial a Pablo por ayudarme a no perder el norte.

A cada una de las personas que han participado en este trabajo, a los grandes profesionales con los que me he cruzado en mi camino y a todos y cada uno de los usuarios que han formado parte desinteresadamente de él.

RESUMEN

Durante los últimos años el número de personas mayores de 65 años ha aumentado considerablemente y se espera que en los próximos años su porcentaje actual llegue a duplicarse. Este envejecimiento progresivo supone un incremento de los costes tanto de intervención como de asistencia sanitaria. En este contexto, las tecnologías de la información y la comunicación (TIC) surgen como una buena oportunidad para el diseño de nuevos servicios de valor añadido que permitan la monitorización de personas mayores disminuyendo los costes y simultáneamente, aumentando la calidad de vida. Este tipo de soluciones promueven la autonomía y el cuidado en el hogar disminuyendo otros factores como la asistencia médica o la tasa de readmisión en hospitales.

La presente tesis doctoral aborda la problemática del envejecimiento progresivo de la población aportando un nuevo método de monitorización de personas que puede ser utilizado para la detección, control y seguimiento de enfermedades. A su vez, el método propuesto permite mejorar la atención e intervención sanitaria apoyando a los cuidadores informales en su labor. Para el diseño del método se ha llevado a cabo un estudio exhaustivo de los principales requisitos que se deben cumplir para el correcto funcionamiento y la aceptación por los usuarios finales. En base a estos requisitos, el método propuesto ha sido validado en un entorno real con cuatro usuarios monitorizados durante tres meses.

El método diseñado y validado en esta tesis está compuesto por tres sub-métodos independientes: método de medida del comportamiento basado en la actividad física, método de modelado para la identificación de rutinas y un método de detección de anomalías. La combinación de estos tres métodos, permite la monitorización automática y adaptativa del comportamiento de los usuarios detectando anomalías de comportamiento.

El método propuesto para la medida del comportamiento se basa en la medida de la actividad física mediante un sensor de aceleración integrado de forma no invasiva en complementos vestibles. Se propone un nuevo estimador de actividad física más eficiente que los existentes hasta el momento y que puede ser implementado en cualquier dispositivo que integre un acelerómetro. Este método ha sido validado de forma individual mediante usuarios con diferentes perfiles de actividad.

Mediante el procesamiento de la actividad física medida, el método de modelado propuesto identifica la rutina del usuario de forma automática y adaptativa. Al contrario que las principales aproximaciones existentes, el método propuesto no requiere de entrenamiento ni de etiquetado. La rutina es construida a partir de las propias medidas históricas y teniendo en cuenta tanto información “intra-diaria” (medidas dentro del propio día) como “inter-diaria” (medidas de los días anteriores). El método propuesto ha sido validado en un entorno real mediante la monitorización de usuarios con diferentes rutinas.

La relación entre la actividad física medida y la rutina identificada es utilizada en el método de detección de anomalías para la identificación de comportamientos anómalos. El objetivo del método de detección es la identificación de medidas de actividad no esperadas según la rutina de la persona monitorizada. Mediante técnicas de lógica difusa se construye de forma adaptativa para cada instante temporal una función de valoración mediante la que se realiza la evaluación del comportamiento. La evaluación del comportamiento consiste en la puntuación de la actividad medida con un valor correspondiente a su grado de anomalía. Una etapa final de filtrado permite la personalización del método para la detección de comportamientos relevantes en función de la persona monitorizada. El método propuesto ha sido analizado en un entorno real mediante la monitorización de usuarios con diferentes perfiles y es capaz de identificar de forma automática tramos temporales de comportamiento anómalo.

Los métodos propuestos han sido validados en un entorno real mediante un sensor de aceleración no invasivo integrado en un reloj con apariencia deportiva. La validación ha tenido lugar de forma individual para cada uno de los métodos y de forma global mediante la monitorización de cuatro personas mayores durante tres meses. La validación se ha realizado en base a los requisitos funcionales y no funcionales identificados: fiabilidad, invasividad, autonomía, adaptabilidad y automatismo. A su vez, se ha llevado a cabo la aplicación del método propuesto para la monitorización de un usuario con enfermedades crónicas bajo supervisión de expertos médicos.

El método propuesto permite la detección automática y adaptativa de anomalías de comportamiento a partir de la medida de la actividad física. Presenta una alta tasa de

precisión y sensibilidad llevando a cabo la detección mediante un sensor de aceleración de forma no invasiva y autónoma para el usuario.

Palabras clave: actividad física, comportamiento anómalo, método adaptativo, método automático, monitorización, personas mayores, rutina, sensor no invasivo, monitorización de personas

ABSTRACT

In recent years the number of people over 65 has increased considerably and it is expected to double in the coming years. This progressive ageing causes an increase in costs both in intervention and healthcare. In this context, information and communication technologies (ICTs) are emerging as a good opportunity for the design of new value-added services for monitoring elderly people as well as decreasing costs and increasing the quality of life. These kinds of solutions promote autonomy and home care as well as the decrease in other factors such as health care assistance or hospital readmission rate.

This thesis addresses the problem of progressive ageing proposing a new method of user monitoring that can be used for the detection, control and tracking of diseases. The proposed method can further improve health care assistance and intervention supporting the informal caregivers work. This work identifies the main requirements needed to get a fully functional solution that provides a broad acceptance by end-users. The method proposed in this work has been validated in a real environment with four users monitored for three months.

The method designed and validated in this thesis is composed of three separate sub-methods: an activity measurement method based on physical activity, a behaviour modelling method for identifying routines and an anomaly detection method. The combination of these methods creates a novel automatic and adaptive monitoring solution to detect anomalous human behaviour.

The proposed activity measurement method is based on the measurement of physical activity using a wearable accelerometer. This method is based on a new and more efficient physical activity estimator than the existing so far. It can be implemented on any device that integrates an accelerometer. It has been validated with users with different activity profiles.

By processing the measured physical activity, the proposed modelling method identifies the user routine automatically and adaptively. Unlike the main existing approaches, the proposed method requires no training or labelling. The routine is built from historical measures taking into account both "intra-daily" measures (measures within the same day) and "inter-daily" measures (measures of previous days). The proposed method has been validated in a real environment by monitoring users with different routines.

The relationship between physical activity and the calculated routine is analysed in the anomaly detection method to detect anomalous behaviour. The term anomalous behaviour refers to the detection of unexpected levels of activity according to the expected user routine. A valuation function is built adaptively for each activity sample using fuzzy logic techniques. This valuation function gives each activity sample a score corresponding to the anomaly degree. The anomaly degree is calculated depending on the relationship between the routine and the measured activity. A final stage of filtering allows the customization of the method to detect only the relevant anomalous behaviour. The proposed method has been tested in a real environment by monitoring users with different profiles and it is able to identify automatically temporal sections of anomalous behaviour.

The proposed methods have been validated in a real environment using a non-invasive accelerometer integrated in a device with a watch-appearance. Each of these methods has been individually validated and besides all of them have been validated as a whole by monitoring four older people for three months. The validation was based on the identified functional and non-functional requirements: reliability, invasiveness, autonomy, adaptability and automation. The proposed method has been further validated with a chronic patient under medical supervision.

The proposed method automatically and adaptively detects anomalous behaviour by measuring physical activity. It has a high rate of accuracy and sensitivity and it is based on the use of a non-invasive and autonomous acceleration sensor.

Keywords: physical activity, anomalous behaviour detection, adaptive method, automatic method, monitoring, elderly, routine, non-invasive sensor, people monitoring

PRÓLOGO

Esta tesis es el resultado de más de seis años de investigación dentro de la unidad 4U – Tecnologías para el usuario (4U) del área I+D+i de Fundación CTIC – Centro tecnológico (CTIC). El objetivo de este grupo de investigación es trabajar en el desarrollo y adaptación de las nuevas tecnologías para conseguir crear soluciones centradas en el usuario y que sean a la vez usables, útiles y universales (4U: *user-centered, usable, useful and universal*). Desde mi incorporación a CTIC como Ingeniero I+D+i en 2008 he trabajado en proyectos de diversas temáticas destacando tres tópicos principales: computación vestible (*wearable computing*) y redes de sensores, inteligencia ambiental y soluciones TIC para la salud y el bienestar (*eHealth*). También he colaborado en proyectos dentro del campo industrial como especialista en técnicas de procesamiento de señal.

En paralelo con mi actividad profesional en CTIC, he colaborado durante casi tres años (2008 - 2011) con el área de Arquitectura y Tecnología de Computadores (ATC) del Departamento de Informática de la Universidad de Oviedo. Dentro de este grupo de trabajo, realicé en 2009 el proyecto fin de carrera: “Valoración de soldaduras mediante el análisis de las variables que influyen en la unión de los materiales” y en 2010 el trabajo fin de máster: “Valoración de la calidad de procesos usando técnicas de procesamiento de la señal” centrados en la aplicación de técnicas de procesamiento de señal para la valoración de procesos industriales.

Con anterioridad a mi incorporación a CTIC, colaboré durante más de un año con el Área de Genética y Reproducción Animal del Servicio Regional de Investigación y Desarrollo Agroalimentario del Principado de Asturias (SERIDA) al amparo de un convenio de colaboración de prácticas de empresa suscrito con la Universidad de Oviedo. Dentro de este convenio, comencé mi actividad investigadora en la aplicación de técnicas matemáticas de análisis para la resolución de problemas y desarrollo de algoritmos.

La combinación de mi actividad actual en CTIC junto con mi experiencia en el área ATC y en el centro de biotecnología animal de SERIDA me ha permitido especializarme en técnicas de procesamiento de variables en diferentes ámbitos y con diferentes objetivos. A su vez, poseo experiencia en el diseño e implementación de redes electrónicas de sensores en aplicaciones de inteligencia ambiental y en el desarrollo y coordinación de pruebas de usabilidad y valoración tecnológica con usuarios.

En el año 2011, a partir de los conocimientos adquiridos durante mi colaboración con el área ATC de la Universidad de Oviedo y mi experiencia en tecnologías centradas en el usuario dentro del área de I+D+i en CTIC surge bajo la supervisión de Víctor Manuel Peláez (responsable de la unidad de investigación 4U) la idea de la que nace esta tesis. El objetivo inicial fue la investigación en la aplicación de las técnicas utilizadas para la valoración de procesos industriales en el campo del comportamiento humano.

Esta tesis se enmarca dentro de un acuerdo de colaboración entre el director de la misma: Gabriel Díaz Orueta, profesor contratado doctor de la Escuela Técnica Superior de Ingenieros Industriales (ETSII) en la Universidad Nacional de Educación a Distancia (UNED); y el co-director: Eduardo Álvarez Álvarez, profesor ayudante doctor de la Escuela Politécnica de Mieres en la Universidad de Oviedo y miembro del equipo directivo de CTIC desde 2005 hasta 2013.

Los resultados contenidos en este trabajo son fruto de varios proyectos en los que he colaborado durante mi trayectoria profesional dentro de CTIC y que se resumen a continuación:

- ELF@Home – “Elderly self-care based on self-check of health conditions and self-fitness at home” (2013-2016). Proyecto desarrollado en el marco de la quinta convocatoria del programa europeo Ambient Assisted Living Joint. Referencia AAL-2012-5-123.
- ELDERCISES – “Sistema basado en la monitorización de la actividad física para la propuesta de ejercicios personalizados en el hogar” (2014-2015). Este proyecto se enmarca dentro del “Programa Asturias” de ayudas a Centros Tecnológicos y Centros de Apoyo a la Innovación Tecnológica en el Principado de Asturias. Referencia CT14-04-2-05.
- TEIXOIS – “Tecnología y Experiencia para generar oportunidades en la Industria. Sistema de detección de cambios y anomalías en la rutina física de las personas para su aplicación en el campo de la eSalud” (2013). Este proyecto se enmarca dentro del “Programa Asturias” de ayudas a Centros Tecnológicos y

Centros de Apoyo a la Innovación Tecnológica en el Principado de Asturias.
Referencia CT13-22.

- SILHOUETTE – “Using novel information and communication technologies for the support of elderly's active participation in the information society” (2011 – 2013). Proyecto desarrollado en el marco del mini-programa Europeo CREATOR, cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER) dentro de la iniciativa europea Interreg IVC.
- cReaTic – “Laboratorio de creatividad e innovación tecnológica” (2008 - 2011). Sistema creado por CTIC para la estimulación de la creatividad e innovación tecnológica dirigido a jóvenes con talento. Financiado por distintas empresas y organismos nacionales e internacionales.
- CETICA - “La ciudad eco-tecno-lógica: industrialización en base acero para un hábitat urbano más sostenible” (2007 – 2010). Programa CENIT – Ingenio 2010. Referencia CEN-20072012.
- ATLAS – “Aplicación de tecnologías asistenciales” (2009 - 2010). Proyecto englobado dentro de la iniciativa "e-inclusión colectivos" de la estrategia e-Asturias. Promovido y ejecutado por Fundación CTIC y financiado por la Consejería de Administraciones Públicas y Portavoz del Gobierno del Principado de Asturias y Telecable S. A. U.

Todos estos proyectos han sido llevados a cabo dentro de un entorno multidisciplinar y de gran especialización tecnológica. En varios de los proyectos como SILHOUETTE o ELF@Home se ha contado con la participación de investigadores extranjeros que han permitido dotar a este trabajo de una perspectiva internacional posicionándolo dentro del contexto en el que actualmente nos encontramos.

Además de los proyectos anteriormente mencionados, he participado en la elaboración y coordinación de propuestas de proyectos regionales, nacionales y europeos. Especialmente, a nivel europeo he participado en el programa de proyectos “Ambient Assisted Living Joint Programme” focalizado en el envejecimiento activo. También he participado como revisor de artículos científicos en “IEEE Transactions on Industrial Informatics” y como evaluador en la convocatoria “PRTS 2013 - Programme de Recherche Transnationale en Santé” promovida en colaboración por la agencia nacional francesa de investigación (French national research agency, ANR) y el ministerio de asuntos sociales y salud (Ministry of social affairs and health, DGOS). A su vez, en los últimos años he colaborado en varios talleres, seminarios y charlas tanto como asistente como ponente.

El conjunto global de la tesis que se presenta en este documento es la suma de los esfuerzos realizados en todos los proyectos de investigación en los que me he involucrado. A lo largo de mi actividad profesional he identificado claramente la necesidad de investigar en nuevos métodos de monitorización de usuarios que cumplan los requisitos necesarios para garantizar su aceptación por personas mayores y que simultáneamente cubran las funcionalidades requeridas. La utilidad de la detección de anomalías de comportamiento para el seguimiento, control y detección de enfermedades ha destacado en varios de los proyectos y propuestas de proyecto en las que he trabajado. Con esta tesis he pretendido dar respuesta a esa necesidad dando un paso más y aportando un nuevo método de detección de anomalías basado en la medida de la actividad física y que de acuerdo a los requisitos necesarios funciona de forma adaptativa, automática, autónoma y no invasiva.

En los próximos años, preveo continuar trabajando en este ámbito continuando con la investigación iniciada en esta tesis y diseminando todos los resultados que he obtenido en la misma.

ÍNDICE

Agradecimientos	i
Resumen	iii
Abstract.....	vii
Prólogo.....	ix
Índice.....	xiii
Índice de figuras.....	xix
Índice de tablas.....	xxiii
CAPÍTULO 1. Introducción	1
1.1. Motivación.....	1
1.2. Objetivos	5
1.3. Hipótesis.....	6
1.4. Organización del documento.....	8
CAPÍTULO 2. Visión general	11
2.1. Introducción	11
2.2. Contexto demográfico.....	12
2.3. Métodos de un sistema de monitorización.....	14
2.3.1. Introducción.....	14
2.3.2. Método de medida.....	14
2.3.3. Método de modelado	16

2.3.4. Método de detección.....	17
2.3.5. Resumen	19
2.4. Arquitectura para la monitorización de personas	19
2.5. Requisitos no funcionales de un sistema de monitorización.....	20
2.5.1. Introducción.....	20
2.5.2. Requisitos de usuario	20
2.5.3. Requisitos técnicos	25
2.5.4. Resumen	26
2.6. Conclusiones.....	26
CAPÍTULO 3. Sistemas de monitorización de personas mayores	29
3.1. Introducción	29
3.2. Estado del arte.....	31
3.2.1. Trabajos de investigación.....	31
3.2.2. Servicios de teleasistencia	34
3.2.3. Proyectos de investigación.....	35
3.3. Análisis del estado del arte.....	37
3.4. Conclusiones.....	40
CAPÍTULO 4. Método de medida: actividad física	41
4.1. Introducción	41
4.2. Estado del arte: medida del comportamiento.....	43
4.3. Actividad física	50
4.3.1. Definición e importancia de la actividad física	50
4.3.2. Métodos para la medida de la actividad física	51
4.4. Diseño del método de medida	55
4.4.1. Selección de una magnitud (RF-3).....	55
4.4.2. Sensor de medida (RF-1)	56
4.4.3. Método de procesado (RF-2).....	58
4.4.4. Validación del método propuesto.....	67
4.4.5. Resumen	73

4.5. Conclusiones	74
CAPÍTULO 5. Método de modelado: identificación de rutina	77
5.1. Introducción	77
5.2. Estado del arte: modelado del comportamiento	79
5.2.1. Introducción.....	79
5.2.2. Método de acondicionamiento.....	80
5.2.3. Método de procesado del comportamiento	87
5.3. Diseño del método de modelado	93
5.3.1. Modelo esperado (RF-6).....	93
5.3.2. Método de acondicionamiento (RF-4).....	94
5.3.3. Método de procesado (RF-5)	107
5.3.4. Validación del método propuesto	111
5.3.5. Resumen.....	114
5.4. Conclusiones	117
CAPÍTULO 6. Método de detección: anomalías de comportamiento	119
6.1. Introducción	119
6.2. Estado del arte: detección de anomalías de comportamiento	121
6.2.1. Introducción.....	121
6.2.2. Métodos no supervisados	123
6.2.3. Métodos supervisados.....	126
6.2.4. Métodos semi-supervisados.....	128
6.2.5. Resumen.....	129
6.3. Diseño del método de detección de anomalías	131
6.3.1. Anomalías detectadas (RF-9).....	131
6.3.2. Método de acondicionamiento (RF-7).....	132
6.3.3. Método de evaluación (RF-8)	132
6.3.4. Validación del método propuesto	145
6.3.5. Resumen.....	151
6.4. Conclusiones	153

CAPÍTULO 7. Validación del método propuesto	155
7.1. Introducción	155
7.2. Diseño del experimento	156
7.2.1. Arquitectura de monitorización.....	156
7.2.2. Usuarios involucrados.....	159
7.2.3. Procedimiento de experimentación.....	161
7.3. Resultados.....	162
7.3.4. Fiabilidad en la detección.....	162
7.3.5. Invasividad y autonomía en la detección	168
7.4. Validez y aplicabilidad médica de los resultados.....	173
7.5. Discusión.....	178
7.6. Conclusiones.....	179
CAPÍTULO 8. Conclusiones y trabajo futuro.....	181
8.1. Introducción	181
8.2. Sistemas de monitorización de personas mayores.....	182
8.2.1. Requisitos identificados	182
8.2.2. Tendencias existentes	183
8.3. Método de medida: actividad física.....	184
8.4. Método de modelado: identificación de rutina	186
8.5. Método de detección: anomalías de comportamiento.....	188
8.6. Validación global del método propuesto	191
8.7. Trabajo futuro.....	192
8.8. Publicaciones científicas	193
Referencias.....	197
APÉNDICE A. Abreviaturas, glosario y nomenclatura.....	217
A.1. Abreviaturas y siglas.....	217
A.2. Glosario de términos.....	218
A.3. Nomenclatura.....	220

APÉNDICE B. Recolección de requisitos de usuario	223
APÉNDICE C. Sistemas de monitorización de personas mayores	227
C.1. Resultados del proyecto SILHOUETTE	227
C.2. Soluciones comerciales.....	230
C.3. Proyectos de investigación	233
C.4. Medidores de actividad física	237
APÉNDICE D. Estudio del método de medida	241
APÉNDICE E. Estudio del método de modelado	245
E.1. Análisis de las principales técnicas de segmentación.....	245
E.1.1. Introducción	245
E.1.2. Segmentación por regiones	245
E.1.3. Segmentación por detección de bordes	248
E.1.4. Enventanado temporal	249
E.2. Lista de usuarios	251
APÉNDICE F. Estudio del método de detección.....	253
F.1. Estudio de métodos semi-supervisados.....	253
F.1.1. Introducción	253
F.1.2. Ajuste a una distribución normal.....	253
F.1.3. Cálculo de histogramas.....	257
F.1.4. Diagramas de cajas	261
F.1.5. Medidas de similitud	264
F.2. Análisis de métodos de evaluación de anomalías	266
F.2.1. Introducción	266
F.2.2. Uso de la desviación estándar del modelo.....	267
F.2.3. Uso de la desviación estándar de la medida.....	268
F.2.4. Funciones de pertenencia fuzzy	269
F.3. Lista de usuarios.....	272
APÉNDICE G. Procedimiento de experimentación	275

G.1. Introducción	275
G.2. Cuestionario de actividad.....	275
G.3. Consentimiento informado	278
APÉNDICE H. Usuario de ejemplo: José García.....	281
H.1. Introducción	281
H.2. Descripción de la rutina del usuario.....	281
H.3. Ejemplo de aplicación del método de modelado propuesto	282
H.4. Ejemplo de aplicación del método de detección propuesto	286
APÉNDICE I. Detalle de los resultados de la validación	291
I.1. Lista de usuarios.....	291
I.2. Comportamientos anómalos detectados.....	292

ÍNDICE DE FIGURAS

Figura 1.1. Evolución de la población mundial por rango de edad.....	1
Figura 1.2. Gasto en salud en los países OECD entre 2000 y 2012	2
Figura 1.3. Consecuencias del uso de sistemas de asistencia basados en TIC.....	3
Figura 1.4. Diagrama general del sistema de monitorización propuesto.....	4
Figura 1.5. Diagrama de objetivos	5
Figura 2.1. Métodos principales de un sistema de monitorización.....	11
Figura 2.2. Infografía europea sobre el uso de las TIC en sistemas de monitorización .	13
Figura 2.3. Etapas de un sistema de monitorización de personas mayores.....	14
Figura 2.4. Diagrama de bloques del método de medida.....	14
Figura 2.5. Diagrama de bloques del método de modelado	16
Figura 2.6. Diagrama de bloques del método de análisis.....	18
Figura 2.7. Elementos de un sistema de monitorización de personas	19
Figura 3.1. Objetivo principal del Capítulo 3	29
Figura 3.2. Agentes que intervienen en un servicio de teleasistencia.....	34
Figura 3.3. Análisis DAFO de los servicios de monitorización de personas	38
Figura 3.4. Elementos que definen el estado de salud de una persona	39
Figura 4.1. Objetivo principal del Capítulo 4	41
Figura 4.2. Relación entre actividad física, condiciones físicas y salud.....	50
Figura 4.3. Características de los principales métodos para medir la actividad física ...	53
Figura 4.4. Aceleración gravitacional de un cuerpo en reposo.....	59
Figura 4.5. Representación gráfica del cálculo de la SMA.....	62
Figura 4.6. Diagrama de bloques para el cálculo de la SMA.....	63
Figura 4.7. Diagrama de bloques para el cálculo del JIM.....	66
Figura 4.8. Pseudocódigo para el cálculo del JIM.....	67
Figura 4.9. Histograma de la correlación entre el JIM y la SMA	69
Figura 4.10. Gráfico de Bland-Altman entre JIM y SMA media por usuario	70
Figura 4.11. Diferencia estimada entre el valor de JIM y SMA para cada usuario	71

Figura 4.12. Comparación entre la implementación de SMA y JIM.....	73
Figura 4.13. Diagrama de bloques del método de medida propuesto.....	73
Figura 5.1. Objetivo principal del Capítulo 5	78
Figura 5.2. Etapas de un método de modelado del comportamiento	79
Figura 5.3. Características que garantizan la calidad de los datos	80
Figura 5.4. Tamaño de ventana	84
Figura 5.5. Tipos de ventanas que pueden ser utilizadas en la segmentación	85
Figura 5.6. Enventanado con y sin solapamiento	85
Figura 5.7. Método de procesado basado en un modelado teórico.....	88
Figura 5.8. Método de procesado basado en reconocimiento de actividad	88
Figura 5.9. Escenarios de un sistema de reconocimiento de actividad.....	89
Figura 5.10. Método de procesado basado en técnicas de regresión.....	90
Figura 5.11. Ejemplo de modelo de comportamiento esperado (rutina).....	94
Figura 5.12. Método de pre-procesamiento propuesto: técnica de limpieza de datos ...	97
Figura 5.13. Pseudocódigo de la estrategia de eliminación de datos redundantes.....	98
Figura 5.14. Pseudocódigo de la estrategia de detección de datos ausentes.....	99
Figura 5.15. Método de segmentación propuesto: enventanado temporal.....	102
Figura 5.16. Pseudocódigo del método de segmentación propuesto.....	102
Figura 5.17. Método de extracción de características propuesto: cálculo de la media .	103
Figura 5.18. Pseudocódigo del método de extracción de características propuesto	104
Figura 5.19. Método de acondicionamiento propuesto.....	105
Figura 5.20. Pseudocódigo del método de acondicionamiento propuesto	105
Figura 5.21. Ejemplo de variable original y variable acondicionada	106
Figura 5.22. Etapa de clasificación del método de modelado propuesto	108
Figura 5.23. Representación gráfica del método de modelado propuesto	109
Figura 5.24. Método propuesto para el modelado del comportamiento	110
Figura 5.25. Pseudocódigo del método de modelado propuesto	110
Figura 5.26. Correlación entre la rutina calculada y la especificada por los usuarios..	112
Figura 5.27. Relación entre el valor de rutina teórica y el modelo obtenido.....	113
Figura 5.28. Etapas del método de modelado propuesto	116
Figura 6.1. Objetivo principal del Capítulo 6	120
Figura 6.2. Principales métodos para la detección de anomalías	123
Figura 6.3. Método de detección de anomalías no supervisado	123
Figura 6.4. Método de detección de anomalías supervisado	126
Figura 6.5. Método de detección de anomalías semi-supervisado.....	128
Figura 6.6. Ejemplo de anomalías detectadas.....	132
Figura 6.7. Fases del método de detección propuesto.....	134
Figura 6.8. Diagrama de bloques de la fase I: cálculo de diferencia de actividad.....	135

Figura 6.9. Pseudocódigo de la fase I: cálculo de diferencia de actividad	136
Figura 6.10. Pseudocódigo de la fase II: extracción de características estadísticas	137
Figura 6.11. Diagrama de la fase II: extracción de características estadísticas	137
Figura 6.12. Compromiso de umbrales para detección de anomalías	139
Figura 6.13. Función de valoración Fuzzy para evaluación de anomalías.....	141
Figura 6.14. Diagrama de la fase III: detección de comportamiento anómalo	142
Figura 6.15. Pseudocódigo de la fase III: detección de comportamiento anómalo	143
Figura 6.16. Pseudocódigo para la aplicación del método de detección propuesto	145
Figura 6.17. Diagrama de análisis de eventos	148
Figura 6.18. Relación entre diferencia de actividad y valor de anomalía	149
Figura 6.19. Variación del parámetro V en el método de detección.....	150
Figura 6.20. Diagrama del método para la detección de anomalías	151
Figura 7.1. Objetivo principal del Capítulo 7	156
Figura 7.2. Prototipo de arquitectura para la monitorización.....	157
Figura 7.3. Componentes del prototipo de elemento concentrador	158
Figura 7.4. Esquema de la arquitectura del elemento remoto	159
Figura 7.5. Distribución de los comportamientos anómalos detectados.....	165
Figura 7.6. Distribución de anomalías detectadas para el usuario 1.....	166
Figura 7.7. Distribución de anomalías detectadas para el usuario 2.....	166
Figura 7.8. Distribución de anomalías detectadas para el usuario 3.....	167
Figura 7.9. Distribución de anomalías detectadas para el usuario 4.....	167
Figura 7.10. Algoritmo para la detección de tiempo de monitorización.....	168
Figura 7.11. Tiempo de monitorización del usuario 1	169
Figura 7.12. Tiempo de monitorización del usuario 2	170
Figura 7.13. Tiempo de monitorización del usuario 3	170
Figura 7.14. Tiempo de monitorización del usuario 4	170
Figura 7.15. Distribución de fallos de autonomía en la monitorización.....	172
Figura E.1. Aplicación de técnicas de segmentación por regiones.....	246
Figura E.2. Aplicación de técnicas de segmentación por fronteras.....	249
Figura E.3. Aplicación de técnicas de inventariado temporal	250
Figura F.1. Función de densidad de probabilidad de una distribución normal	254
Figura F.2. Ajuste a una distribución normal de las medidas de actividad.....	256
Figura F.3. Cálculo de histograma para detección de anomalías	259
Figura F.4. Detección de anomalías mediante histograma normalizado	261
Figura F.5. Diagrama de cajas para la detección de anomalías	262
Figura F.6. Aplicación de diagramas de cajas para detección de anomalías	263
Figura F.7. Aplicación de la distancia euclídea como medida de similitud	265
Figura F.8. Detección mediante umbral estático.....	267

Figura F.9. Detección mediante umbral basado en la desviación del modelo.....	268
Figura F.10. Detección mediante umbral basado en la desviación de la medida.....	269
Figura F.11. Función de pertenencia triangular	270
Figura F.12. Función de pertenencia trapezoidal.....	270
Figura F.13. Función de pertenencia gamma.....	271
Figura F.14. Función de pertenencia sigmoideal	271
Figura F.15. Función de pertenencia gaussiana	271
Figura G.1. Ejemplo de construcción de actividad teórica	277
Figura H.1. Medidas de actividad física durante $N_r = 5$ días.....	283
Figura H.2. Medidas de actividad física acondicionadas durante $N_r = 5$ días.....	284
Figura H.3. Salida del método de modelado: rutina	285
Figura H.4. Medida de actividad física del Día 6 de monitorización.....	286
Figura H.5. Ejemplo de aplicación del método de detección de anomalías (I).....	287
Figura H.6. Ejemplo de aplicación del método de detección de anomalías (II)	288

ÍNDICE DE TABLAS

Tabla 1.1. Evolución esperada por cada 100 personas en edad laboral.....	2
Tabla 2.1. Requisitos funcionales del método de medida	15
Tabla 2.2. Requisitos funcionales del método de procesamiento	17
Tabla 2.3. Requisitos funcionales del método de análisis.....	18
Tabla 2.4. Resumen de las entrevistas del proyecto SILHOUETTE.....	21
Tabla 2.5. Deseo de las personas mayores a permanecer en su hogar.....	21
Tabla 2.6. Funcionalidades de interés para una solución TIC para personas mayores..	22
Tabla 2.7. Principales requisitos de usuario para un sistema de monitorización	24
Tabla 2.8. Principales requisitos técnicos para un sistema de monitorización.....	25
Tabla 3.1. Tendencias en sistemas de monitorización de personas mayores	32
Tabla 3.2. Tipos de sistemas de monitorización de personas según su objetivo.....	33
Tabla 3.3. Iniciativas europeas relacionadas con servicios de monitorización.....	36
Tabla 4.1. Medida del comportamiento: sensores PIR.....	43
Tabla 4.2. Medida del comportamiento: sensores de localización	43
Tabla 4.3. Medida del comportamiento: vídeo y/o micrófonos.....	44
Tabla 4.4. Medida del comportamiento: detectores de consumo eléctrico	44
Tabla 4.5. Medida del comportamiento: etiquetas RFID	44
Tabla 4.6. Medida del comportamiento: detectores magnéticos	45
Tabla 4.7. Medida del comportamiento: sensores de presión.....	45
Tabla 4.8. Medida del comportamiento: medidores de cantidad de movimiento	45
Tabla 4.9. Tipos principales de sensores para la adquisición de medidas	46
Tabla 4.10. Principales alternativas de sensores vestibles	48
Tabla 4.11. Métodos estándar de medida de actividad física	52
Tabla 4.12. Métodos objetivos de medida de actividad física.....	52
Tabla 4.13. Métodos subjetivos de medida de actividad física.....	53
Tabla 4.14. Características de un acelerómetro para medir actividad física.....	57
Tabla 4.15. Características acelerométricas para la medida de la actividad física.	61

Tabla 4.16. Consumo eléctrico de varios acelerómetros comerciales.....	65
Tabla 4.17. Usuarios involucrados en la validación del método de medida.....	68
Tabla 4.18. Comparación el cálculo de la SMA y del JIM	72
Tabla 4.19. Resumen de los parámetros implicados en el método de medida	73
Tabla 4.20. Requisitos funcionales del método de medida.....	74
Tabla 5.1. Estrategia de integración de datos	81
Tabla 5.2. Estrategia de limpieza de datos	81
Tabla 5.3. Estrategia de transformación de datos	82
Tabla 5.4. Estrategia de reducción de datos.....	82
Tabla 5.5. Principales técnicas de segmentación	83
Tabla 5.6. Principales tipos de características.....	86
Tabla 5.7. Principales técnicas de regresión para el modelado.....	91
Tabla 5.8. Parámetros del método de segmentación propuesto	100
Tabla 5.9. Usuarios involucrados en la validación del método de modelado	112
Tabla 5.10. Resumen de los parámetros implicados en el método de modelado.....	114
Tabla 5.11. Requisitos funcionales del método de procesamiento	115
Tabla 6.1. Detección de anomalías en el estado del arte	122
Tabla 6.2. Técnicas de agrupamiento: detección basada en pertenencia.....	124
Tabla 6.3. Técnicas de agrupamiento: detección basada en distancia.....	125
Tabla 6.4. Técnicas de agrupamiento: detección basada en densidad	125
Tabla 6.5. Principales técnicas de clasificación de comportamiento	127
Tabla 6.6. Comparación de métodos para la detección de anomalías	130
Tabla 6.7. Usuarios involucrados en la validación del método de detección	145
Tabla 6.8. Resultados en base a los cuestionarios de actividad	148
Tabla 6.9. Resultados del método de detección variando el parámetro V	150
Tabla 6.10. Resumen de los parámetros implicados en el método de detección.....	152
Tabla 6.11. Requisitos funcionales del método de análisis	152
Tabla 7.1. Resumen de las características técnicas del TI eZ430-Chronos.....	157
Tabla 7.2. Datos resumen de usuarios	160
Tabla 7.3. Datos detallados de los usuarios	160
Tabla 7.4. Tabla para la presentación de comportamientos anómalos	163
Tabla 7.5. Comportamientos anómalos detectados y su distribución temporal	164
Tabla 7.6. Tiempos de monitorización para cada usuario	169
Tabla 7.7. Resumen de fallos de autonomía en la monitorización	172
Tabla 7.8. Parámetros del filtro de valores de anomalía	176
Tabla 7.9. Comportamientos anómalos detectados	176
Tabla C.1. Resultados obtenidos en la ejecución del proyecto SILHOUETTE.....	228
Tabla C.2. Soluciones comerciales para la monitorización de personas.....	230

Tabla C.3. Principales proyectos de investigación en monitorización y teleasistencia	233
Tabla C.4. Principales medidores comerciales de actividad física	237
Tabla D.1 Lista de usuarios involucrados en la validación del método de medida	242
Tabla D.2 Resultados de la validación del método de medida.....	243
Tabla E.1. Lista de usuarios y resultados de la validación del método de modelado..	251
Tabla F.1. Lista de usuarios involucrados en la validación del método de detección.	272
Tabla F.2. Comportamientos anómalos identificados por los usuarios	273
Tabla F.3. Resultados del método de detección variando el parámetro V	274
Tabla G.1. Cuestionario de actividad física	276
Tabla G.2. Ejemplo de cuestionario de actividad física	277
Tabla H.1. Comportamientos anómalos identificados en el Día 6	286
Tabla I.1. Lista de usuarios involucrados en la validación global del método.....	291
Tabla I.2. Comportamientos anómalos detectados para el usuario 1	292
Tabla I.3. Comportamientos anómalos detectados para el usuario 2	293
Tabla I.4. Comportamientos anómalos detectados para el usuario 3	294
Tabla I.5. Comportamientos anómalos detectados para el usuario 4	295

CAPÍTULO 1.

INTRODUCCIÓN

1.1. MOTIVACIÓN

El Departamento de Economía de las Naciones Unidas (United Nations, Department of Economic and Social Affairs 2013b) prevé para los próximos años que se doble el porcentaje de población mundial mayor de 60 años. La Figura 1.1, extraída de (United Nations, Department of Economic and Social Affairs 2013a), muestra la distribución de población por rango de edad. En ella se aprecia el aumento esperado de la población mundial mayor de 60 años. De la misma forma se prevé que el ratio de personas mayores frente a personas en edad laboral aumente considerablemente en los próximos años. La proyección esperada puede verse representada en la Tabla 1.1.

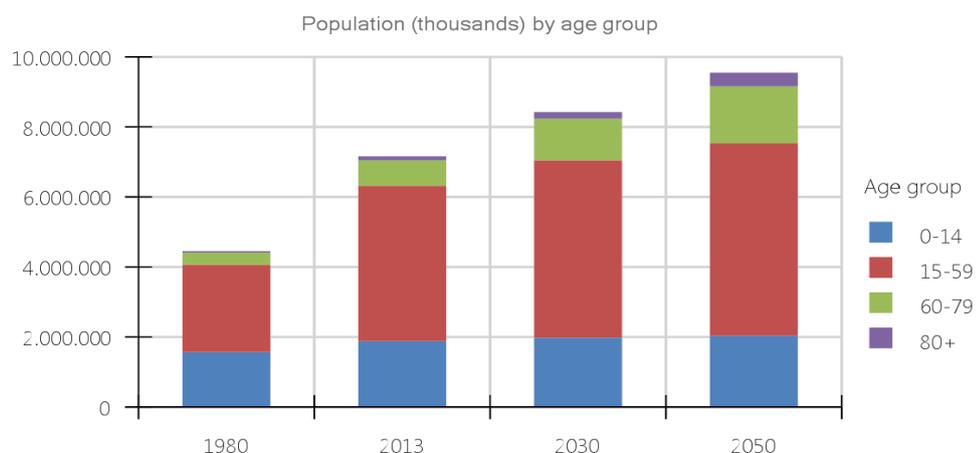


Figura 1.1. Evolución de la población mundial por rango de edad

Tabla 1.1. Evolución esperada por cada 100 personas en edad laboral

	1980	2013	2030	2050
Niños	59,9	39,9	36,2	33,8
Personas mayores	10,2	12,1	17,8	24,7
Total	70,1	52,0	54,0	58,5

La principal consecuencia del envejecimiento progresivo de la población es la necesidad del aumento de la inversión en atención e intervención sanitaria. Este hecho se acentúa con la alta tasa de migración y urbanización actual que hace que aumente el número de personas mayores que viven solas en su hogar y que necesitan de asistencia. Sin embargo, tal y como se muestra en la Figura 1.2 extraída de (Organización para la Cooperación y Desarrollo Económicos 2012), en el actual contexto económico la inversión en intervención y asistencia sanitaria se muestra estancada. Es necesaria la propuesta de nuevos métodos de atención que satisfagan las necesidades de la sociedad y que supongan una solución económicamente sostenible.

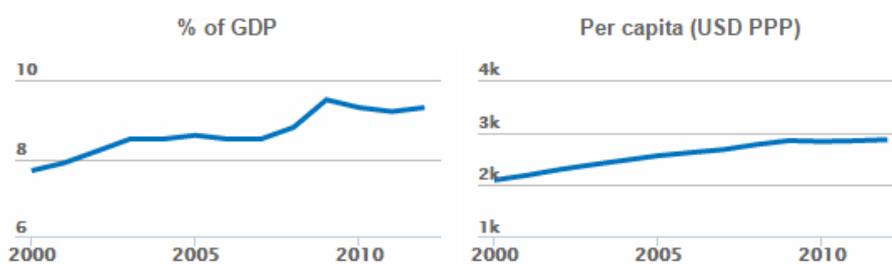


Figura 1.2. Gasto en salud en los países OECD entre 2000 y 2012

Las tecnologías de la información y la comunicación (TIC) suponen una gran oportunidad en el desarrollo de nuevos sistemas de asistencia y han demostrado ampliamente su efectividad en diferentes ámbitos tanto en la detección de riesgos como en el control y seguimiento de enfermedades (Patel and Benzo 2007) (Coye, Haselkorn, and DeMello 2009) (Darkins, Ryan, and Kobb 2008). Los servicios de asistencia basados en TIC (ver Figura 1.3) permiten principalmente prolongar la vida del usuario en su hogar, mejorar su calidad de vida y suponen un método de apoyo para la labor de los cuidadores informales (Aanesen, Lotherington, and Olsen 2011) (Brignell, Wootton, and Gray 2007) (Steele, Min, and Lo 2012). El concepto “cuidador informal” engloba a personas que dedican un importante porcentaje de su vida diaria al cuidado de personas mayores con los que generalmente mantienen un vínculo familiar. En Europa, se estima que un 80% de los cuidados recibidos por personas mayores provienen de cuidadores informales (Courtin, Jemai, and Mossialos 2014).

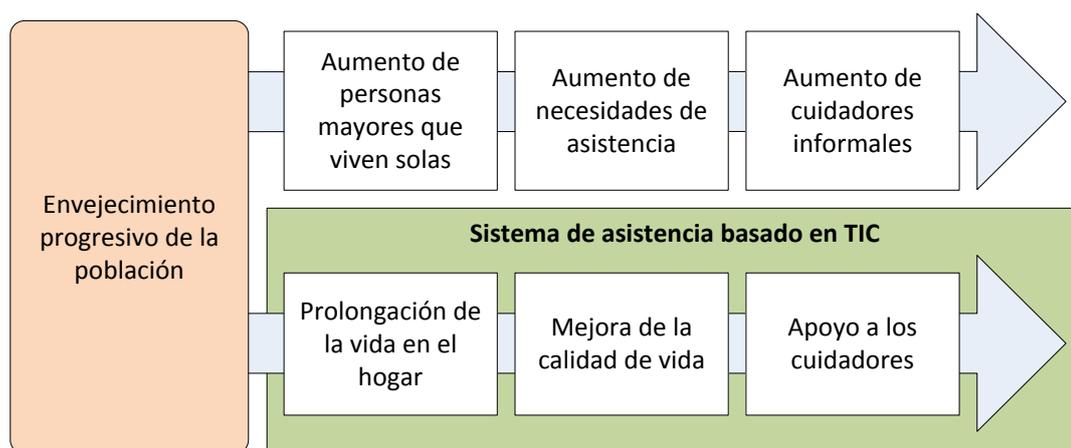


Figura 1.3. Consecuencias del uso de sistemas de asistencia basados en TIC

En la actualidad existen dos tendencias principales en el uso de las TIC para el cuidado de personas mayores: monitorización de variables médicas y el conocido como “botón del pánico”. El último de estos sistemas permite al usuario el envío de alarmas cuando se encuentra en situaciones de peligro para que automáticamente se realice el aviso a sus cuidadores. La monitorización de variables médicas permite el seguimiento del estado de salud del usuario mediante el uso de sensores vestibles. En ambos sistemas, el usuario debe conocer el manejo del sistema. Este hecho supone la principal barrera para su penetración en la sociedad debido a las características y habilidades del usuario (tanto físicas como psicológicas) (Ziefle and Rocker 2010).

En este trabajo se propone un servicio de teleasistencia para personas mayores mediante el diseño de un sistema de monitorización basado en TIC. El objetivo principal es prolongar la permanencia de los usuarios en el hogar, mejorar su calidad de vida y apoyar a los cuidadores informales en su labor. El sistema de monitorización propuesto funcionará de forma automática y transparente (no invasiva) para los usuarios minimizando su interacción y suponiendo una nueva vía de cuidado para los cuidadores informales.

El sistema propuesto se basa en la monitorización de personas mayores mediante la medida de la actividad física (ver esquema general en la Figura 1.4). Esta magnitud, ampliamente utilizada con fines deportivos, es definida como cualquier movimiento del cuerpo humano que se traduce en un gasto energético (Caspersen, Powell, and Christenson 1985) y es un buen indicador del estado de salud de una persona (Vanhees et al. 2005). Hasta el momento, esta magnitud ha estado principalmente relacionada con la monitorización del estilo de vida del usuario midiendo su grado de sedentarismo. En este trabajo, se propone su uso como medida del comportamiento para la detección de patrones no esperados.

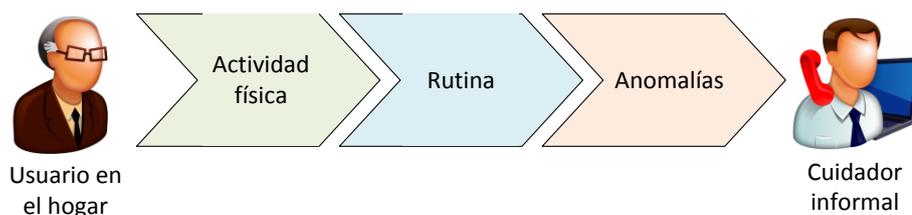


Figura 1.4. Diagrama general del sistema de monitorización propuesto

El sistema propuesto, mediante el procesamiento de la actividad física, identifica la rutina del usuario (comportamiento esperado). Esta rutina es construida de forma automática y adaptativa sin realizar ninguna asunción teórica y basándose únicamente en las medidas de actividad tomadas. El análisis del comportamiento del usuario en un día dado respecto a su rutina es utilizado para la detección de anomalías que pueden ser reportadas a los cuidadores informales para su información y posible intervención. Por detección de anomalías se entiende la detección en una colección de datos de patrones no esperados (Chandola, Banerjee, and Kumar 2009). En el campo del comportamiento humano, se entiende por detección de anomalías a la detección de cambios significantes en el nivel de actividad de un usuario inusuales de acuerdo a su rutina (Lühr, West, and Venkatesh 2007) (Antonakaki, Kosmopoulos, and Perantonis 2009).

Las principales innovaciones de este trabajo frente al estado del arte pueden resumirse en:

- Uso de la actividad física como magnitud de monitorización del estado de salud del usuario. Se propone un novedoso y eficiente método de medida de la actividad física basado en el procesamiento de la aceleración del cuerpo humano.
- Identificación de forma automática y adaptativa de la rutina del usuario. A través de la rutina se define el modelo de comportamiento esperado para el usuario. Esta rutina es construida basada en las propias medidas tomadas de actividad sin realizar ninguna asunción teórica sobre el comportamiento realizado o esperado.
- Detección de anomalías mediante el análisis de la relación entre la medida de actividad tomada y la rutina identificada en el usuario. La detección de anomalías se basa en el uso de técnicas estadísticas y de minería de datos para la identificación de tramos temporales en los cuales la actividad del usuario difiere considerablemente de lo esperado. La detección se realizará de forma automática con el objetivo de informar a los cuidadores informales de posibles problemas y/o trastornos de comportamiento.

1.2. OBJETIVOS

Este trabajo persigue como objetivo principal el diseño y validación de un método automático y adaptativo de monitorización de usuarios que cumpla todos los requisitos necesarios para su uso en personas mayores. Entre estos requisitos destaca principalmente el diseño de un sistema fiable, no invasivo y de alta autonomía. La Figura 1.5 representa de forma gráfica los objetivos principales de este trabajo.

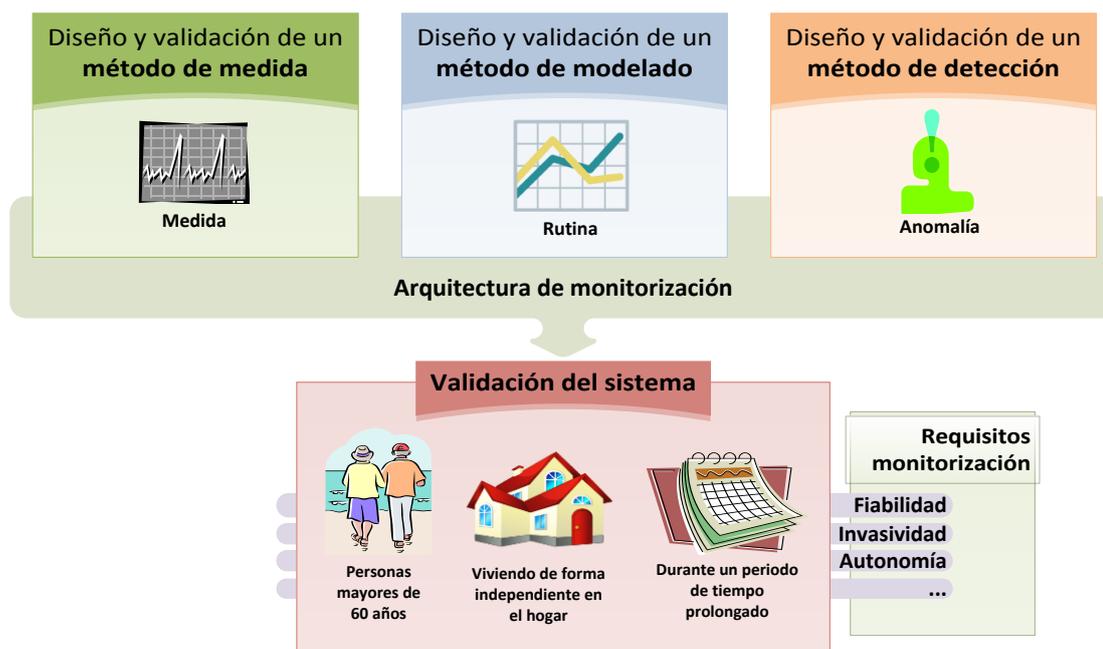


Figura 1.5. Diagrama de objetivos

Como objetivos específicos puede realizarse la siguiente enumeración:

- Estudio, identificación y análisis de los principales requisitos que debe cumplir un sistema de monitorización para su uso por personas mayores.
- Diseño y validación de un método de medida del comportamiento humano basado en la actividad física.
- Diseño y validación de un método de modelado del comportamiento humano que permita la identificación de la rutina del usuario.
- Diseño y validación de un método de detección de anomalías que permita la detección de comportamientos no usuales de acuerdo a la rutina del usuario.

Para la obtención de estos objetivos específicos, será preciso el diseño de una arquitectura de monitorización adecuada para su uso en personas mayores y la implementación de un prototipo que permita la validación del sistema propuesto.

En base a los objetivos anteriormente citados, las principales contribuciones esperadas en este trabajo serán:

- Colección de requisitos que debe cumplir un sistema de monitorización para su aceptación por personas mayores. Esta colección de requisitos servirá como guía para el desarrollo de nuevos sistemas de monitorización de personas mayores.
- Método de medida del comportamiento humano que permita la medida de la actividad física de forma no invasiva. Este método supondrá un avance sobre los métodos ya existentes primando la eficiencia y efectividad del mismo.
- Método de modelado del comportamiento humano que permita la identificación de rutinas de forma automática y adaptativa. El método propuesto detectará un modelo de comportamiento esperado y podrá ser aplicado en otros ámbitos con el objetivo de identificar rutinas o patrones de comportamiento basados en datos históricos.
- Método de detección de anomalías que permita detectar comportamientos anómalos de forma automática y adaptativa. El método propuesto servirá para la detección genérica de anomalías de comportamiento en base a la medida de la actividad física pero podrá ser configurado para la detección de anomalías específicas relacionadas con trastornos o dolencias concretas.

1.3. HIPÓTESIS

Las tecnologías de la información y la comunicación (TIC) pueden ser utilizadas para el desarrollo de un nuevo sistema de monitorización de personas mayores que permita la detección, control y seguimiento de enfermedades. La efectividad y eficiencia en el uso de las TIC para la monitorización de personas mayores ha sido probada en varios trabajos en los que se demuestra cómo su uso permite mejorar la calidad de vida de las personas y apoyar a la labor de los cuidadores y expertos sanitarios. Sin embargo, los sistemas de monitorización existentes en la actualidad presentan una baja tasa de penetración en la sociedad ya que no cumplen los requisitos necesarios para su uso por personas mayores.

Es necesario identificar los requisitos tanto funcionales como no funcionales que debe cumplir un sistema de monitorización para que sea aceptado por los usuarios prestando especial atención a las necesidades de las personas mayores. A nivel funcional un sistema de monitorización debe estar formado por un método de medida, un método de modelado y un método de detección. La funcionalidad de cada uno de

estos métodos depende del objetivo de la monitorización y debe funcionar de forma automática y adaptativa. Dentro de los requisitos no funcionales destacan principalmente los requisitos técnicos y los requisitos de usuario. A nivel técnico, el sistema de monitorización debe ser fiable y seguro presentando una alta tasa de detección y una baja tasa de falsas alarmas. A nivel de usuario, se debe buscar un sistema no invasivo y autónomo que no requiera de su interacción para su correcto funcionamiento.

Dentro de los sistemas de monitorización actuales destaca la monitorización del comportamiento. La medida del comportamiento del usuario permite la detección de situaciones que pueden estar relacionadas con trastornos como hiperactividad o insomnio. En el caso de personas mayores, la monitorización del comportamiento permite detectar el conocido como síndrome de fragilidad y enfermedades como Alzheimer o demencia.

En este trabajo se propone el diseño de un nuevo método de monitorización de personas mayores basado en el uso de la actividad física para la medida del comportamiento. El objetivo de este nuevo método de monitorización es la detección de anomalías de comportamiento de forma automática y adaptativa. Por anomalía de comportamiento se entiende la detección de niveles de actividad no esperados de acuerdo a la rutina del usuario. Este nuevo método de monitorización está formado por tres métodos independientes: método de medida de la actividad física, método de modelado para la identificación de rutinas y método de detección de anomalías.

El diseño de estos métodos estará guiado por los requisitos identificados para su uso en personas mayores que guiarán su validación. Los métodos propuestos serán validados en primer lugar de forma independiente y a continuación de forma global. La validación individual de cada método propuesto tendrá lugar en un escenario real y con usuarios de diferentes edades para garantizar la experimentación con alta variabilidad de actividad.

La validación global del sistema propuesto tendrá lugar con al menos cuatro personas mayores que serán monitorizadas en condiciones de vida libre y en un entorno real durante al menos 90 días. Se evaluará la correcta detección de anomalías de comportamiento (precisión y sensibilidad) y se analizará la consecución de los requisitos identificados en las primeras etapas del trabajo. Para ello se estudiará el tiempo de monitorización y se analizarán los posibles errores detectados. La validación del sistema tendrá lugar mediante la implementación de una arquitectura que garantice la monitorización y detección de anomalías de comportamiento en un entorno real de funcionamiento.

La hipótesis de este trabajo podría resumirse en el siguiente párrafo:

Un método de monitorización de comportamiento basado en la medida de la actividad física puede detectar anomalías de comportamiento de forma automática, adaptativa y cumpliendo todos los requisitos necesarios para su uso en personas mayores.

1.4. ORGANIZACIÓN DEL DOCUMENTO

En este primer capítulo (Capítulo 1) se describe la motivación de la que nace esta investigación, los objetivos del mismo y la hipótesis de partida. Esta hipótesis será validada durante el resto de capítulos que componen el trabajo.

El Capítulo 2 presenta una visión general del trabajo atendiendo al contexto demográfico que motiva esta investigación y a los tres métodos que componen el sistema: método de medida, método de modelado y método de detección. A su vez se muestran los principales elementos que componen una arquitectura para la monitorización de personas y se presenta la colección de requisitos que el sistema de monitorización deberá cumplir para su uso por personas mayores. Estos requisitos se han obtenido mediante el análisis de los principales trabajos encontrados en el estado del arte y el cuestionario recogido en el Apéndice B. Esta colección de requisitos guiará el diseño, implementación y validación del trabajo.

El Capítulo 3 describe y analiza el estado del arte en sistemas de monitorización para personas mayores. Se presentan y analizan las principales tendencias existentes a nivel de investigación, de servicios y de proyectos. El análisis DAFO (debilidades, amenazas, fortalezas y oportunidades) de los sistemas de monitorización existentes pone de manifiesto la oportunidad que supone el uso de las TIC para la monitorización de personas mayores. En este capítulo se destacan también las ventajas de un sistema de monitorización basado en el análisis del comportamiento. Este capítulo se complementa con la información contenida en el Apéndice C.

El Capítulo 4 se centra en el diseño del método de medida de comportamiento. Se realiza un análisis del estado del arte en cuanto a sensores y magnitudes existentes para la medida destacando el uso de acelerómetros vestibles y de la actividad física. Se propone y se valida el uso de un nuevo método de medida que cumple todos los requisitos impuestos para la monitorización de personas mayores. En el Apéndice D se recogen detalles acerca de la etapa de experimentación y validación llevada a cabo en este capítulo.

En el Capítulo 5 se describen las principales alternativas existentes para el modelado del comportamiento humano y la identificación de rutinas. Se identifican dos etapas diferenciadas en los métodos existentes (acondicionamiento y procesado) y se analizan las principales aproximaciones encontradas en el estado del arte para cada una de estas fases. En la parte final del capítulo se propone y valida un nuevo método de modelado basado en técnicas de regresión. El Apéndice E complementa este capítulo mostrando parte de los resultados científicos obtenidos y detalles de la etapa de experimentación y validación.

El Capítulo 6 se centra en el diseño de un método que permita la detección de anomalías de comportamiento a partir de los métodos de medida y modelado propuestos en los capítulos anteriores. En este capítulo se estudian y analizan tres alternativas principales para la detección de anomalías: métodos no supervisados, métodos supervisados y métodos semi-supervisados. Estos últimos son los seleccionados en este trabajo. Dentro de los métodos semi-supervisados se estudian las principales técnicas existentes y se diseña y valida un método dividido en tres fases para la detección. El método propuesto se basa en el cálculo de los residuos (cálculo de la diferencia) entre la medida de la actividad y la rutina identificada (fase I) para la detección de anomalías a través de una función de pertenencia Fuzzy (fase III) creada de forma dinámica en base a las características estadísticas de las variables (fase II). El Apéndice F contiene parte de los resultados obtenidos en este capítulo y detalles de la experimentación y validación del método propuesto.

El Capítulo 7 se centra en la etapa de experimentación en la que se realiza la validación de los métodos propuestos con cuatro usuarios en condiciones de vida libre. Para llevar a cabo esta validación se utiliza una arquitectura compuesta por un elemento de medida (dispositivo vestible), un elemento concentrador (ordenador portátil) y un elemento remoto (colección de servicios web). Esta etapa de experimentación ha tenido lugar durante 90 días en los que los usuarios fueron monitorizados de forma no invasiva mediante un sensor de aceleración integrado en un reloj. La validación tiene lugar en base a los requisitos identificados en el Capítulo 2. Para ello se estudia la precisión en la detección, el tiempo de monitorización y los errores detectados durante la experimentación. En la fase final de este capítulo se presenta el sistema a una asociación médica y se prolonga la validación con un usuario que presenta varias enfermedades crónicas entre las que destaca un riesgo cardiovascular alto. Los detalles acerca de los resultados obtenidos en la validación global del método pueden consultarse en el Apéndice I.

El Capítulo 8 describe las conclusiones extraídas tras la realización de este trabajo y las posibles líneas futuras en las que se prevé continuar la investigación. Dentro de este capítulo se destacan las principales aportaciones a nivel científico de este trabajo y las publicaciones científicas derivadas.

El Apéndice A recoge las principales siglas y abreviaturas utilizadas a lo largo de este trabajo. Además, se presenta una colección de términos comúnmente usados con el objetivo de clarificar su significado. Por último, se dedica una sección a la presentación de la nomenclatura utilizada en la definición de los métodos propuestos.

Además de los apéndices ya mencionados, el Apéndice G y el Apéndice H contiene información genérica acerca de las etapas de experimentación y validación llevadas a cabo en este trabajo. En el caso del Apéndice H se describe un usuario de ejemplo, cuya actividad y rutina son utilizadas a lo largo del trabajo para la descripción y validación de los métodos y técnicas propuestas. El Apéndice G detalla el procedimiento seguido para la validación.

CAPÍTULO 2.

VISIÓN GENERAL

2.1. INTRODUCCIÓN

Este capítulo presenta una visión general del problema que se pretende resolver en este trabajo y que ya fue brevemente introducido en el capítulo anterior. Partiendo del contexto demográfico actual (descrito en la sección 2.2) caracterizado por el envejecimiento progresivo de la población, se pone de manifiesto la necesidad y oportunidad de invertir en nuevos sistemas de asistencia basados en TIC. Este tipo de soluciones permiten prolongar la vida del usuario en el hogar y ayudan a los cuidadores informales al desarrollo de su actividad.

Siguiendo estas ideas, en este trabajo se propone el desarrollo de un sistema de monitorización de personas mayores basado en la medida de la actividad física para la detección de anomalías de comportamiento (comportamientos anómalos). Los tres métodos principales de los que se compone un sistema de monitorización se presentan en la Figura 2.1 y son descritos en la sección 0: método de medida, método de modelado y método de detección. En este capítulo, se pone especial énfasis en los requisitos funcionales que debe cubrir cada método y que guiarán el diseño del sistema.

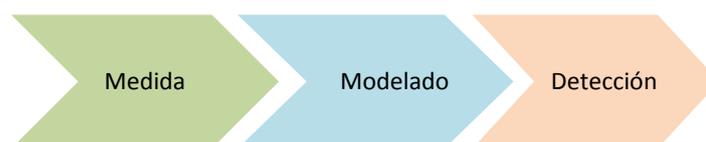


Figura 2.1. Métodos principales de un sistema de monitorización

La implementación de los métodos propuestos requiere el diseño de una arquitectura de monitorización que es introducida en la sección 2.4 y que está compuesta por tres elementos principales: elemento de medida, elemento concentrador y elemento remoto. Los métodos de medida, modelado y detección propuestos serán implementados sobre esta arquitectura para su estudio, análisis y validación.

El diseño general del sistema propuesto debe tener en cuenta simultáneamente los requisitos funcionales de cada método junto con los requisitos técnicos y los requisitos de usuario necesarios para garantizar su aceptación. Estos requisitos no funcionales (técnicos y de usuario) son identificados en la sección 2.5. Este último tipo de requisitos cobran gran importancia y serán el factor diferenciador de este trabajo que es guiado por los requisitos que un sistema de monitorización debe cumplir para su uso en personas mayores. Estos requisitos se han identificado mediante el análisis del estado del arte y encuestas a más de 80 personas.

2.2. CONTEXTO DEMOGRÁFICO

Según las previsiones del Departamento de Economía de las Naciones Unidas (United Nations, Department of Economic and Social Affairs 2013b), la población mundial mayor de 65 años aumentará en los próximos años de un 11,7% del total en 2013 a 21,1% en 2015. Los motivos de este incremento están directamente relacionados con el decremento de la fertilidad y con el incremento de la esperanza de vida (Rechel et al. 2013). En el caso de Europa (Economist Intelligence Unit 2011), se espera un incremento de las personas mayores de un 16% en el año 2000 al 24% en 2030.

Dentro de la población actual de personas mayores un 40% de la población mayor de 60 años vive sola y de forma independiente (United Nations, Department of Economic and Social Affairs 2013b). Esta tasa se encuentra en incremento y varía en función de la zona de estudio. Mientras que en el norte de Europa, un 90% de la población vive de forma independiente (cifra similar a Estados Unidos (Alwan, Wiley, and Nobel 2007)), en el sur (Grecia, Italia, España y Portugal) solamente un 40% vive en su hogar. Estas diferencias están condicionadas principalmente por aspectos culturales y sociales.

A medida que la población envejece, la necesidad de invertir en atención sanitaria aumenta y se pone en relieve la necesidad de crear soluciones tecnológicas que permitan mejorar la atención reduciendo costes (Aanesen et al. 2011). El uso de sistemas de monitorización basados en el uso de las TIC y de tecnologías vestibles permite cubrir esa necesidad (Karunanithi 2007) (Shah et al. 2013).

La Figura 2.2 muestra una infografía desarrollada por la comisión europea que representa el contexto demográfico actual y cómo el uso de las TIC permite mejorar la calidad de vida y ofrecer nuevas oportunidades de negocio y desarrollo.

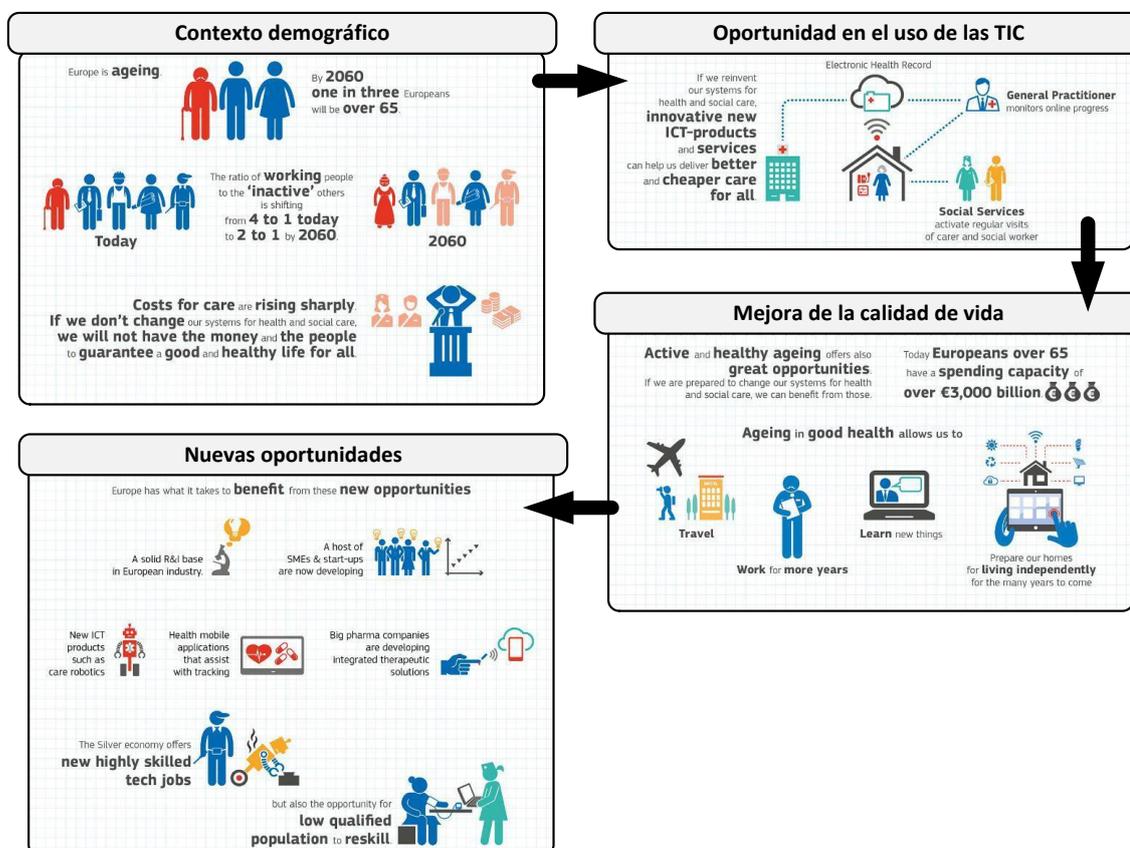


Figura 2.2. Infografía europea sobre el uso de las TIC en sistemas de monitorización

Partiendo del contexto demográfico actual, este trabajo propone un sistema de asistencia basado en la monitorización remota de personas mayores que permitirá reducir los costes en la atención y mejorar la eficiencia en la intervención. El sistema remoto de monitorización podrá enviar alertas a los cuidadores informales y les apoyará en el desarrollo de su actividad.

2.3. MÉTODOS DE UN SISTEMA DE MONITORIZACIÓN

2.3.1. INTRODUCCIÓN

Un sistema de monitorización de personas puede dividirse en los métodos funcionales que se describen en la Figura 2.3: método de medida, método de modelado y método de detección.

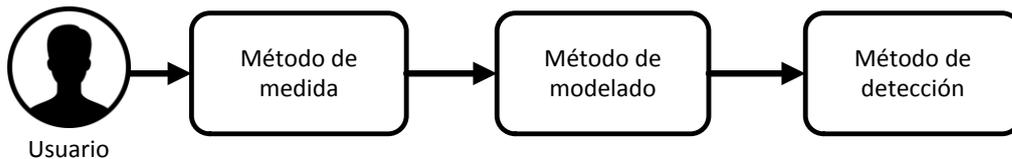


Figura 2.3. Etapas de un sistema de monitorización de personas mayores

Esta división parte de las guías dadas en (Heil, Brage, and Rothney 2012) para el diseño de sistemas de monitorización y se utiliza a nivel general en todos los sistemas de inteligencia ambiental como el descrito en (Aztiria, Izaguirre, and Augusto 2010) y (Simon Brownsell and Bradley 2011) o los sistemas propuestos en (Estevez-Tapiador, Garcia-Teodoro, and Diaz-Verdejo 2004) y (Xie et al. 2011) para la detección de anomalías.

2.3.2. MÉTODO DE MEDIDA

El método de medida es el encargado de adquirir datos del usuario y puede dividirse en los bloques que se muestran en la Figura 2.4:

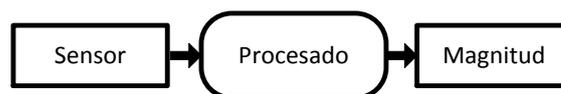


Figura 2.4. Diagrama de bloques del método de medida

- *Sensor*. El método de medida se basa en la adquisición (activa o pasiva) de información de un sensor. Los sensores ambientales y los sensores vestibles son las dos principales alternativas (Atallah and Yang 2009):
 - o *Sensores ambientales*. Sensores integrados en el entorno. Dependiendo del grado de integración puede conseguirse una transparencia total para el usuario.
 - o *Sensores vestibles*. Sensores portados por el usuario. Son más precisos pero más invasivos que los sensores ambientales al tener que ser vestidos por el usuario en todo momento.

- *Procesado*. Las medidas adquiridas en el sensor son procesadas para obtener la magnitud de entrada al método de modelado. La lógica del bloque de procesado transformará las medidas adquiridas del sensor en la magnitud utilizada para la monitorización.

Así, el diseño de un método de medida abarca el diseño de un método de procesado de los datos adquiridos por un sensor para la obtención de una magnitud de medida. Los datos adquiridos variarán en función del objetivo del sistema de monitorización (variables médicas, comportamiento, localización...). En este trabajo el método de medida será diseñado para adquirir y procesar una variable que permita caracterizar el comportamiento.

Los requisitos funcionales que debe cumplir el método de medida son resumidos en la Tabla 2.1. Deben seleccionarse tanto sensores como magnitudes estándar y ampliamente aceptadas por la comunidad para garantizar el diseño de un método de medida fiable (Heil et al. 2012). El método de procesado es el punto clave del método de medida. A través de él se realiza la transformación de los datos de salida del sensor a una magnitud útil para su procesamiento en el siguiente método funcional (método de modelado).

Tabla 2.1. Requisitos funcionales del método de medida

Referencia	Requisito	Descripción
RF-1	Sensor de medida	Sensor utilizado para la monitorización. El método de medida se basará en el uso de un sensor para la adquisición de información del usuario
RF-2	Método de procesado	Se deberá implementar un método que realice un procesado de las medidas adquiridas para obtener la magnitud de medida
RF-3	Magnitud de medida	La magnitud seleccionada para realizar la monitorización del usuario será la salida del método de medida y deberá ser una magnitud estándar

2.3.3. MÉTODO DE MODELADO

Cuando el objetivo de la monitorización es la detección de anomalías, el principal objetivo del método de modelado es la construcción de un modelo esperado de comportamiento (rutina) con el que realizar el análisis en la siguiente etapa del sistema.

La Figura 2.5 muestra el diagrama de bloques del método de modelado. Este método recibe como entrada la magnitud de salida del método de medida. En este trabajo, el objetivo del método de modelado es la identificación de la rutina del usuario a través de la medida utilizada para caracterizar su comportamiento. El método de modelado a diseñar está compuesto por un método de acondicionamiento y un método de procesado:

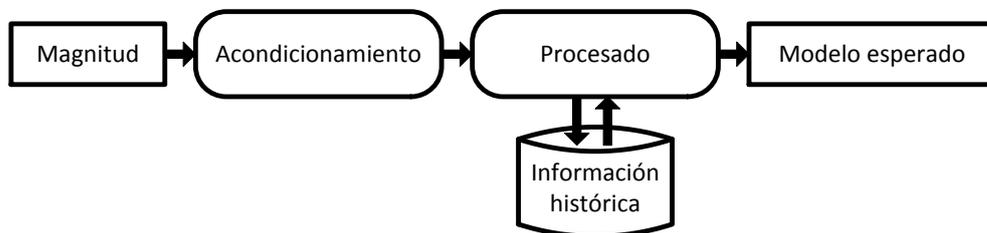


Figura 2.5. Diagrama de bloques del método de modelado

- *Acondicionamiento.* La magnitud de salida del método de medida es la entrada de este método y será acondicionada mediante un bloque de acondicionamiento. En este bloque se aplicarán técnicas de pre-procesamiento de la señal para la eliminación de ruido y/o errores y garantizar la calidad de la información de entrada al bloque de procesado y al método de análisis.
- *Procesado.* A través de diferentes técnicas, el principal objetivo del bloque de procesado es extraer conocimiento de la información adquirida para su análisis. El algoritmo seleccionado para el procesado debe ser dinámico y fácilmente ajustable a los cambios que puedan tener lugar. Para ello, se propone el uso de información histórica que permitirá que el modelado se realice de forma adaptativa. Las técnicas utilizadas para el procesado variarán en función de la naturaleza de la medida y del conocimiento que se desee analizar.
- *Modelo esperado.* La salida del bloque de modelado será una variable que represente el modelo de comportamiento esperado y que será posteriormente analizado en el método de análisis para la detección de anomalías.

Tabla 2.2. Requisitos funcionales del método de procesamiento

Referencia	Requisito	Descripción
RF-4	Método de acondicionamiento	El método de procesamiento deberá implementar un método para el acondicionamiento de la variable eliminando ruido y posibles errores
RF-5	Método de procesado	El núcleo principal del método de procesamiento será el método de modelado en el que residirá toda la lógica del método para la obtención de un modelo esperado
RF-6	Modelo esperado	El modelo esperado será una variable que represente la rutina del usuario y que será la entrada al método de análisis

La Tabla 2.2 muestra los requisitos que deben cumplirse en el diseño del método de modelado. El método de procesado es el método principal del método de modelado ya que tras la eliminación de posibles errores de la medida en la etapa de acondicionamiento, será el encargado de extraer las características relevantes de la magnitud utilizada construyendo el modelo esperado (Heil et al. 2012).

2.3.4. MÉTODO DE DETECCIÓN

Dado un método de medida del comportamiento y un método para el modelado del mismo, el objetivo del método de detección será diferente en función del propósito del sistema dentro del cual se realiza. En este trabajo, el objetivo del método de detección será la identificación de anomalías de comportamiento. Por anomalías se entienden cambios de comportamiento no esperados.

Tras realizar el procesamiento de las medidas, el método de detección se encarga de evaluar la relación entre el modelo esperado (salida del método de modelado) y la magnitud de medida (salida del método de medida). La Figura 2.6 muestra el diagrama de bloques en los que puede dividirse el método de análisis. La magnitud de medida y el modelo esperado son las entradas al método:

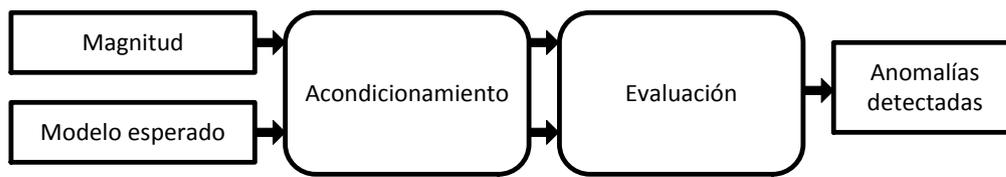


Figura 2.6. Diagrama de bloques del método de análisis

- *Acondicionamiento.* Con el mismo objetivo que en el método de procesamiento, se aplicarán técnicas de pre-procesamiento sobre la magnitud de medida y sobre el modelo esperado para permitir su análisis y comparación.
- *Evaluación.* Las técnicas empleadas en el bloque de evaluación tendrán como objetivo la comparación de la magnitud medida y el modelo esperado para la detección de anomalías. El método utilizado para la evaluación será el encargado de detectar periodos de comportamiento anómalo en la actividad del usuario.
- *Anomalías detectadas.* Las anomalías detectadas se identificarán a través del instante temporal en el que tienen lugar, su duración y un valor que represente su grado de importancia.

El resumen de los requisitos que se deben cumplir en el diseño del método de detección se muestra en la Tabla 2.3. El método utilizado para la evaluación de la variable procesada (salida del método de procesamiento) es el núcleo principal del diseño de este método funcional.

Tabla 2.3. Requisitos funcionales del método de análisis

Referencia	Requisito	Descripción
RF-7	Método de acondicionamiento	El método de acondicionamiento realizará transformaciones en las variables de entrada para permitir su comparación. Este método se basará en el uso de técnicas de pre-procesamiento
RF-8	Método de evaluación	El método de evaluación deberá analizar la relación entre la magnitud medida y el modelo esperado
RF-9	Anomalías detectadas	La salida del método de detección deberá ser una variable con la que se identifique la aparición de anomalías señalando instante temporal, duración y magnitud

2.3.5. RESUMEN

Para cada uno de los métodos que componen un sistema de monitorización se han identificado los requisitos funcionales que deben cumplir. Cada uno de estos requisitos se corresponde con un bloque funcional que será necesario diseñar y evaluar en este trabajo.

Para conseguir satisfacer el conjunto global de requisitos funcionales (RF-1, RF-2, RF-3, RF-4, RF-5, RF-6, RF-7, RF-8 y RF-9) se ha de diseñar un sistema de monitorización que sea automático y adaptativo:

- *Automático.* Los métodos propuestos deben funcionar sin configuración e interacción humana (tanto de los propios usuarios como de expertos técnicos).
- *Adaptativo.* Los métodos propuestos deben adaptarse a las particularidades de cada usuario y a su posible variación temporal.

2.4. ARQUITECTURA PARA LA MONITORIZACIÓN DE PERSONAS

Para la validación de los métodos funcionales anteriormente descritos es necesario el diseño de una arquitectura para la monitorización de personas. Esta arquitectura ha sido diseñada previamente en (Carús and Díaz 2014) para su uso en personas mayores partiendo de todos los requisitos encontrados en el estado del arte y en la propia experiencia con usuarios.

Un sistema de monitorización de personas está compuesto de forma genérica por los tres elementos principales representados en la Figura 2.7: elemento de medida, elemento concentrador y elemento procesador. Cada uno de estos elementos debe ser diseñado junto con el sistema de comunicación utilizado para la transmisión de información entre ellos:

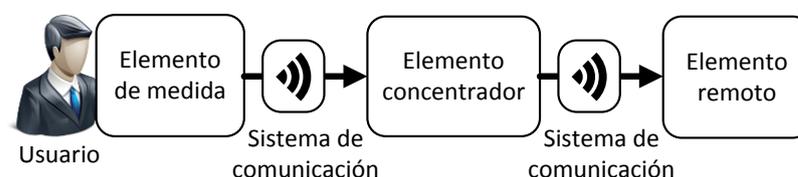


Figura 2.7. Elementos de un sistema de monitorización de personas

- *Elemento de medida.* El elemento de medida es el dispositivo encargado de adquirir la información del usuario. A nivel funcional deberá implementar el método de medida y un sistema de comunicación que permita el envío de la magnitud medida al elemento concentrador de información.
- *Elemento concentrador.* El elemento concentrador actúa como una pasarela de datos. Su principal objetivo es recibir la información adquirida por el elemento de medida y hacérsela llegar al elemento remoto. A nivel funcional deberá implementar los sistemas y protocolos de comunicación que le permitan la recepción de información del elemento de medida y el envío de ésta al elemento remoto.
- *Elemento remoto.* Este elemento recibe la información adquirida por el elemento de medida a través del elemento concentrador y realiza su procesamiento y análisis. A nivel funcional deberá implementar los métodos de modelado y detección.

2.5. REQUISITOS NO FUNCIONALES DE UN SISTEMA DE MONITORIZACIÓN

2.5.1. INTRODUCCIÓN

Esta sección presenta los principales requisitos no funcionales (requisitos de usuario y requisitos técnicos) que debe cumplir un sistema de monitorización de personas mayores de acuerdo a la información recolectada en diversos estudios. Estos requisitos deben ser seguidos para el diseño de los tres métodos funcionales descritos en la sección 0: método de medida, método de modelado y método de detección.

2.5.2. REQUISITOS DE USUARIO

Para estudiar los requisitos que un sistema de monitorización basado en el uso de TIC debe satisfacer de acuerdo a las necesidades de las personas mayores, se ha llevado a cabo un estudio mediante la entrevista a 83 personas mayores residentes en Asturias. La edad media de los entrevistados fue de 72 años con un rango de edad entre los 61 y los 91 años. Estas entrevistas se enmarcaron dentro del proyecto europeo SILHOETTE y además de en Asturias tuvieron lugar en el resto de regiones participantes en el proyecto: Häme (Finlandia), Lorraine (Francia), Norte de Hungría, Valle Sabbia (Italia), Västerbotten (Suecia) y Wielkopolska (Polonia). La Tabla 2.4 muestra el resumen de las entrevistas realizadas y en el Apéndice B puede consultarse el cuestionario utilizado.

Tabla 2.4. Resumen de las entrevistas del proyecto SILHOUETTE

	Asturias 	Häme 	Lorraine 	Norte Hungria 	Valle Sabbia 	Väster- botten 	Wielko- polska 
Número de encuestados	83	100	100	100	46	37	53
Media de edad	72	77	76	69	74	78	73
Rango de edad	64-91	65-96	62-89	65-86	62-90	65-92	62-84

En la primera parte de las entrevistas, se preguntó por el deseo de los entrevistados a permanecer en su hogar suponiendo la existencia de una pérdida de autonomía y planteando la posibilidad del uso de una solución TIC. Tal y como muestra la Tabla 2.5 se validó una de las principales motivaciones de este trabajo: prolongar la vida de las personas mayores en su hogar.

En el caso de Asturias, un 88% de las personas mayores encuestadas desean seguir residiendo en su hogar y en el caso de que se llevara a cabo una pérdida de autonomía preferirían adoptar soluciones de asistencia antes que moverse a un lugar especializado. Sin embargo, solamente un 19% de esas personas se plantearían el uso de soluciones TIC para prolongar la vida en su hogar.

Tabla 2.5. Deseo de las personas mayores a permanecer en su hogar

¿Viviría en su hogar..	Asturias 	Häme 	Lorraine 	Norte Hungria 	Valle Sabbia 	Väster- botten 	Wielko- polska 
...ahora?	88%	90%	92%	98%	89%	89%	91%
...ante una pérdida de autonomía?	34%	78%	43%	-	32%	43%	60%
...con una solución TIC?	19%	33%	25%	-	15%	8%	9%

Durante la última parte de las entrevistas y dentro de un ambiente distendido se intentó identificar cuáles son los principales pros y contras para que un porcentaje tan bajo de los encuestados estén dispuestos a valorar el uso de una solución TIC para prolongar su vida en el hogar. Los encuestados conciben las soluciones TIC como soluciones complicadas de manejar y que requieren de conocimiento técnico para su funcionamiento. Esto hace que en general exista una desconfianza ante este tipo de soluciones pese a que son conscientes de su posible utilidad y sienten curiosidad por su uso.

Planteando la posibilidad de utilizar una solución TIC, se preguntó a los encuestados sobre cuál, según su criterio, debería ser la finalidad de esa solución para que supusiera una mejora en su calidad de vida. De entre todas las opciones planteadas, la Tabla 2.6 muestra por cada región cuáles fueron las tres soluciones principales.

En el caso de Asturias, un 51% de las personas encuestadas estarían interesadas en una solución que aumente su comunicación con familiares y expertos en salud, un 48% en una solución TIC de teleasistencia y un 42% en una solución que aumente y garantice su seguridad. A nivel global las soluciones de teleasistencia, seguridad y comunicación destacan como las principales funcionalidades de interés para las personas mayores.

A partir del estudio realizado, se determina que para conseguir una buena involucración de los usuarios es necesario mostrar la utilidad del sistema con el que se pretende trabajar. Simultáneamente una solución TIC para personas mayores debe cumplir una serie de características para garantizar su involucración.

Tabla 2.6. Funcionalidades de interés para una solución TIC para personas mayores

¿Solución TIC para...	Asturias 	Häme 	Lorraine 	Norte Hungría 	Valle Sabbia 	Västerbotten 	Wielkopolska 
...ocio?			<input checked="" type="checkbox"/>				<input checked="" type="checkbox"/>
teleasistencia?	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
...seguridad?	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
comunicación?	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>	
...tareas del hogar?			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		

Las principales características que los usuarios buscan en un sensor para su monitorización según varios estudios ((Gokalp and Clarke 2013), (Murphy 2009), (Steele et al. 2009), (Ziefle and Rucker 2010), (Wu et al. 2012), (Fessl, Rivera-Pelayo, and Müller 2011), (Pantelopoulos and Bourbakis 2010), (Avci and Bosch 2010), (Aztiria et al. 2010)) y las encuestas llevadas a cabo en este trabajo (SILHOUETTE Partnership 2013a) se describen en la Tabla 2.7. En esta relación de requisitos también se tienen en cuenta las conclusiones extraídas durante el proyecto ATLAS (“Aplicación de tecnologías asistenciales”) y que se relatan en (Fundación CTIC 2010). En la ejecución de este proyecto se destacó la necesidad de dispositivos para la monitorización de personas mayores de uso intuitivo, ligeros, autónomos y con un sistema sencillo de carga de batería. Todo este conjunto de características fijan los requisitos de usuario que un sistema de monitorización de personas mayores debe cumplir.

Los requisitos impuestos por los usuarios (RU) pueden resumirse en dos características principales:

- *No invasividad.* (RU-1, RU-2 y RU-3). El sistema de monitorización debe ser embebido en dispositivos que los usuarios suelen utilizar diariamente de forma que el sistema no suponga una molestia para la realización de las actividades de la vida diaria. Esta aproximación resuelve también los problemas que puede acarrear el uso del sistema para personas con problemas cognitivos.
- *Autonomía.* (RU-4, RU-5, RU-6, RU-7 y RU-8). El sistema de monitorización debe ser autónomo requiriendo la mínima interacción posible del usuario y con una memoria de almacenamiento y tiempo de vida de la batería dimensionada adecuadamente según las necesidades de la aplicación.

Una de las principales medidas para valorar si un sistema de monitorización cumple los requisitos necesarios para su uso por personas mayores, es el cálculo de la tasa de uso o tiempo de monitorización. Según el estudio realizado en (Ledger 2014), la mayoría de los sistemas de monitorización presentan una alta tasa de uso al comienzo pero al cabo de 5 meses, el tiempo de monitorización decae un 25% respecto a la inicial. Para resolver este problema y conseguir una tasa de uso alta que se prolongue en el tiempo es necesario trabajar en los requisitos mencionados en la Tabla 2.7.

Tabla 2.7. Principales requisitos de usuario para un sistema de monitorización

Referencia	Requisito	Descripción
RU-1	Comodidad	Los dispositivos deben ser cómodos tanto si su uso es constante como si se utiliza para medidas puntuales. Dispositivos incómodos para el usuario terminarán siendo no utilizados
RU-2	Tamaño	Los dispositivos para facilitar su manejo y minimizar la intrusión deberían ser pequeños y ligeros
RU-3	Invisibilidad	El dispositivo no debe llamar la atención a los visitantes. Dispositivos grandes y/o difíciles de no ser vistos suponen inconvenientes para los usuarios
RU-4	Facilidad de uso	El dispositivo utilizado para la monitorización debe ser sencillo de manejar por el usuario siendo éste capaz de realizar las tareas mínimas de mantenimiento (encender, apagar, reiniciar)
RU-5	Automático	La configuración del sistema, mantenimiento y posibles actualizaciones debería poder realizarse telemáticamente y sin requerir interacción del usuario. De esta forma se minimiza la interacción humana y posibles errores
RU-6	Memoria	Los dispositivos utilizados deben ser capaces de almacenar datos en el caso de que estos no puedan ser enviados a algún elemento externo
RU-7	Alimentación	El tiempo de vida de la batería de los dispositivos debe ser maximizada y el método de alimentación debe requerir la mínima interacción por parte del usuario
RU-8	Cobertura	El rango de cobertura del sistema de monitorización debe ser lo suficientemente amplio para no interferir en las actividades de la vida diaria del usuario y poder ser utilizado tanto dentro como fuera del hogar

2.5.3. REQUISITOS TÉCNICOS

Los principales requisitos técnicos (RT) que debe cumplir un sistema de monitorización se resumen en la Tabla 2.8.

Tabla 2.8. Principales requisitos técnicos para un sistema de monitorización

Referencia	Requisito	Descripción
RT-1	Fiabilidad	El sistema ha de presentar una alta precisión y una alta sensibilidad. La inexactitud en las medidas hace que la aceptabilidad del sistema por parte del usuario y cuidadores decrezca
RT-2	Seguridad	La seguridad en la comunicación e intercambio de información debe estar implícita en cualquier sistema de monitorización de personas. En este tipo de sistemas se trabaja generalmente con información altamente sensible en los que el requisito de seguridad es crucial

A nivel técnico (RT-1 y RT-2) el sistema de monitorización ha de ser fiable y seguro. Estas dos características son fundamentales para que el sistema sea aceptado tanto por parte de los usuarios como por parte de los cuidadores y permitirá validar la eficiencia y efectividad (Xie et al. 2011) (Aztiria et al. 2010).

La fiabilidad de un sistema de monitorización está directamente relacionada con el objetivo del método de detección. Generalmente, el objetivo es llevar a cabo la detección de algún evento de interés y la fiabilidad se estudia a partir de las medidas de precisión y sensibilidad. Para que un sistema de monitorización sea fiable ha de presentar una alta precisión y una alta sensibilidad (*recall*). La precisión hace referencia a la exactitud en la detección y representa la fracción de eventos detectados correctamente respecto al conjunto total de los detectados. Esta medida permite conocer cuál es el porcentaje de eventos detectados que se corresponden con eventos reales. La sensibilidad, también conocida como exhaustividad, cobertura o *recall*, representa del total de eventos ocurridos el porcentaje que el sistema de monitorización es capaz de detectar. Existe una medida conocida como *F-score* que permite representar con un solo valor la fiabilidad del sistema en base a las medidas de precisión y sensibilidad. La tasa de detección y la tasa de falsos positivos (eventos detectados y no ocurridos) también son útiles para la medida de la fiabilidad del

sistema. Un sistema fiable será aquel que simultáneamente presente una alta tasa de detección junto con una baja tasa de falsos positivos.

El otro requisito técnico, relacionado con la seguridad, implica la existencia de mecanismos que permitan encriptar la información y garantizar su no acceso por terceros. Para ello, todas las interfaces de comunicación presentes en el sistema de monitorización deben poseer mecanismos de encriptado e identificación que garanticen la confidencialidad en el intercambio de información y su protección frente a posibles ataques externos. En la mayoría de los sistemas de monitorización se trabaja con información altamente sensible.

2.5.4. RESUMEN

El estudio realizado de los requisitos no funcionales de un método de monitorización de personas cuyo objetivo es la detección de anomalías pone de manifiesto la necesidad de realizar un diseño buscando un método fiable, seguro, no invasivo y autónomo. Estas cuatro características son perseguidas durante la ejecución de este trabajo y son cubiertas por el sistema de asistencia propuesto.

2.6. CONCLUSIONES

El envejecimiento progresivo de la población hace que aumenten los costes en atención e intervención sanitaria. La investigación en el uso de las tecnologías de la información (TIC) para la monitorización de personas mayores permite reducir estos costes a la vez que se aumenta la seguridad y se obtienen mayores beneficios en la atención.

Un sistema de monitorización de personas mayores debe cumplir una serie de requisitos tanto funcionales como no funcionales para su correcto funcionamiento y para garantizar su aceptación por los usuarios. Dentro de los requisitos no funcionales, en este trabajo se ha identificado la necesidad de que el sistema sea seguro y fiable (requisitos técnicos) y que funcione de forma autónoma y no invasiva para el usuario (requisitos de usuario). El diseño de un sistema de monitorización siguiendo los requisitos establecidos supone un avance respecto a las principales soluciones existentes en la actualidad consiguiendo reducir las barreras en su uso y garantizando la funcionalidad necesaria para la monitorización.

En este trabajo se identifican tres métodos principales en un sistema de monitorización de personas: método de medida, método de modelado y método de detección. Para cada uno de estos métodos se han identificado los requisitos funcionales necesarios que

deben cumplir y que deberán guiar su diseño. Los métodos y requisitos identificados son comunes para cualquier sistema de monitorización de personas y podrán ser utilizados en futuros trabajos guiando el diseño de nuevos e innovadores sistemas de monitorización.

Para la implementación de los métodos identificados, se propone una arquitectura de monitorización de personas basada en la existencia de tres elementos principales: elemento de medida, elemento concentrador y elemento remoto. Aunque esta arquitectura es utilizada en varias aproximaciones existentes en el estado del arte, en este trabajo se identifica claramente la funcionalidad que debe cubrir cada elemento. La arquitectura debe cubrir simultáneamente los requisitos funcionales de los métodos de monitorización (medida, modelado y detección) y los requisitos técnicos y de usuario que garanticen la validez y aceptación del sistema.

CAPÍTULO 3.

SISTEMAS DE MONITORIZACIÓN DE PERSONAS MAYORES

3.1. INTRODUCCIÓN

Este capítulo describe los sistemas de monitorización de personas mayores existentes en la actualidad y recoge las principales tendencias existentes en el estado del arte. Para ello se ha llevado a cabo un minucioso estudio identificando trabajos y proyectos de investigación en el ámbito y analizando los servicios de teleasistencia disponibles comercialmente. La Figura 3.1 muestra de forma gráfica el objetivo de este capítulo.



Figura 3.1. Objetivo principal del Capítulo 3

La sección 3.2 recoge las principales tendencias de monitorización diferenciando trabajos de investigación (sección 3.2.1), servicios de teleasistencia (sección 3.2.2) y proyectos (sección 3.2.3). Dentro de los trabajos de investigación (principalmente artículos científicos) se identifican los tipos de sistemas de monitorización existentes según su objetivo (detección, evaluación, petición de servicios, motivación, educación o inclusión) y las principales medidas utilizadas en ellos (medida del comportamiento, medidas fisiológicas y medidas subjetivas). Dentro de los servicios de teleasistencia tanto a nivel europeo como a nivel mundial se identifican en la sección 3.2.2 las principales alternativas existentes destacando el servicio conocido popularmente como “botón del pánico” y los servicios de recolección de variables médicas para la evaluación del estado de salud.

Existen varias iniciativas a nivel europeo en las que se engloban diversos proyectos de investigación con el objetivo de diseñar nuevos métodos de monitorización de personas mayores. Las principales iniciativas y proyectos identificados se recogen en la sección 3.2.3. La mayoría de estos proyectos trabajan en el entorno del hogar creando nuevas formas de monitorización de usuarios que permitan evaluar su estado de salud con el objetivo de intervenir en el caso de que sea necesario.

En la última parte de este capítulo (sección 3.3) se realiza el análisis del estado de arte recolectado detectando las debilidades, amenazas, fortalezas y oportunidades (análisis DAFO) de los sistemas de monitorización existentes. La identificación de las principales debilidades y amenazas que sufren los sistemas de monitorización en la actualidad permite trabajar sobre sus fortalezas y oportunidades intentando mitigar su efecto. Un nuevo sistema de monitorización que parta de las fortalezas, aprovechando las oportunidades, mitigando las debilidades y luchando contra las amenazas tendrá garantizada la penetración en la sociedad y su eficiencia y efectividad.

A partir del análisis realizado, se detecta la necesidad de invertir esfuerzos en el diseño de sistemas de monitorización basados en el análisis del comportamiento mediante el uso de las TIC, que estén específicamente diseñados para personas mayores y cuyo objetivo sea mejorar su calidad de vida.

3.2. ESTADO DEL ARTE

3.2.1. TRABAJOS DE INVESTIGACIÓN

En la Tabla 3.1 se describen las principales medidas utilizadas para la adquisición de datos en sistemas de monitorización de personas mayores (Gokalp and Clarke 2013): medidas de comportamiento, medidas fisiológicas y medidas subjetivas. Para cada tipo de medida se muestra su objetivo, el tipo de sensores que se pueden emplear y se referencian trabajos dentro del estado del arte que las utilizan.

Además de los trabajos referenciados en la Tabla 3.1, existen también aproximaciones que combinan diferentes tipos de medidas como (Medjahed and Istrate 2009), (Papadopoulos, Crump, and Wilson 2010) o (Raad and Yang 2008). En estos trabajos se realiza una monitorización de las actividades realizadas por el usuario a partir de medidas de actividad (principalmente sensores de presencia) y de medidas fisiológicas (sensores de salud).

Como aplicaciones finales de los trabajos identificados en el estado del arte, deben destacarse los sistemas orientados a la detección y prevención de caídas, los sistemas de localización de usuarios (tanto en interiores como en exteriores) y los sistemas de monitorización del estado de salud (a través de medidas fisiológicas). De este tipo de aplicaciones, desde el punto de vista del usuario final y sus cuidadores, en (Wild et al. 2008) se destacan aquellos sistemas de monitorización que permitan prolongar la vida en su propio hogar, detectar problemas cognitivos incipientes y compartir información con expertos sanitarios.

Los sistemas de monitorización de personas mayores pueden ser divididos, tal y como se muestra en la Tabla 3.2, en seis categorías en función del objetivo de evaluación para el cual se realiza la adquisición de datos (Ludwig et al. 2012). Las dos últimas categorías descritas en la Tabla 3.2 (sistemas de educación y consulta e inclusión social) son consideradas en este trabajo sistemas de monitorización por la existencia de un canal de comunicación entre personas y cuidadores/expertos y por su objetivo de facilitar y mejorar la calidad de vida de las personas. Sin embargo, no son en realidad sistemas que permitan monitorizar al usuario sino que facilitan su comunicación con el exterior y el conocimiento de sus condiciones sanitarias.

Los sistemas de detección de condiciones adversas cobran gran importancia ya que permiten el seguimiento y control de enfermedades facilitando la intervención e incrementando la efectividad en la asistencia.

Tabla 3.1. Tendencias en sistemas de monitorización de personas mayores

Medida	Objetivo
Medidas de comportamiento	Detección de problemas relacionados con el comportamiento: rutinas, movilidad, sedentarismo, caídas
Estudio de las actividades realizadas por el usuario	Sensores
	Sensores de presencia, sistemas de videovigilancia, detectores de consumo de electricidad, sensores RFID, sensores de localización, acelerómetros, podómetros
	Referencias
	(Fouquet and Franco 2010) (Fontecha and Hervás 2011) (Kasteren, Englebienne, and Kröse 2010) (Hajj Chehade et al. 2012) (Fleury, Vacher, and Noury 2010) (Virone 2009) (Doukas and Maglogiannis 2008)
Medida	Objetivo
Medidas fisiológicas	Detección de síntomas asociados a determinadas enfermedades o que puedan ser precursores de dolencias
Monitorización de variables médicas para la detección de situaciones críticas	Sensores
	Presión sanguínea, pulso cardiaco, electrocardiograma, flujo respiratorio, peso
	Referencias
	(Chen 2011) (Heise and Skubic 2010) (Huo et al. 2009) (Bourouis, Feham, and Bouchachia 2011) (Navarro, Lawrence, and Lim 2009) (Lee et al. 2010) (Singh, Lubecke, and Boric-Lubecke 2011) (Yamana et al. 2011)
Medida	Objetivo
Medidas subjetivas	Recopilación de datos subjetivos de los usuarios: uso de elementos, opinión, percepción
El propio usuario aporta información al sistema sobre sus condiciones	Sensores
	Cuestionarios, diarios de actividad, patrones de sueño
	Referencias
	(Wu et al. 2012) (Englert et al. 2010) (Li et al. 2010) (Paladini et al. 2010) (Craig et al. 2003)

Tabla 3.2. Tipos de sistemas de monitorización de personas según su objetivo

Objetivo	Descripción	Tipos
Detección de condiciones adversas	Servicios que pretenden proteger al usuario de posibles situaciones de peligro	<ul style="list-style-type: none"> - Pulsadores de emergencia - Detección de comportamientos anómalos - Detección de caídas - Detección de situaciones de peligro
Evaluación del estado de salud	Servicios que evalúan el estado de salud de la persona. Este estado de salud es accesible por sus cuidadores formales (médicos) e informales (familiares)	<ul style="list-style-type: none"> - Detección de enfermedades y/o trastornos - Monitorización de enfermedades y/o trastornos - Monitorización de intervenciones terapéuticas
Petición de servicios	Servicios que pueden ser utilizados por el usuario para requerir ayuda en las actividades de la vida diaria (ADL)	<ul style="list-style-type: none"> - Petición de compra - Petición de comida - Petición de limpieza
Motivación y recordatorios	Servicios de motivación y recordatorio para conseguir la correcta adherencia del usuario a un tratamiento	<ul style="list-style-type: none"> - Recordatorios de medicación - Calendarios - Servicios de mensajería - Motivación en el uso de teleasistencia
Educación y consulta	Servicios que pretenden educar al usuario en el manejo y conocimiento de sus propias enfermedades	<ul style="list-style-type: none"> - Servicios de teleconferencia - Dispositivos de consejo - Asistencia virtual - Recopiladores de información
Inclusión social	Servicios que pretenden garantizar la inclusión social del usuario. Este tipo de iniciativas entran dentro de los requisitos de una "vida saludable"	<ul style="list-style-type: none"> - Servicios de videoconferencia - Redes sociales específicas para personas mayores

3.2.2. SERVICIOS DE TELEASISTENCIA

Los servicios de teleasistencia pueden ser definidos como servicios tecnológicos que permiten mejorar la calidad de vida de personas dependientes prolongando el tiempo de permanencia en su hogar y en el caso de personas mayores, mejorando el proceso de envejecimiento (Alwan et al. 2007). Esta mejora debe ser entendida a nivel de salud (mejora del estado de salud de las personas) y a nivel de calidad de vida (mejora de la percepción y satisfacción del usuario con su entorno). Los usuarios de estos servicios tecnológicos son tanto las personas mayores como sus cuidadores (familiares y/o expertos sanitarios).

La Figura 3.2 muestra los principales agentes que intervienen en un servicio de teleasistencia y la relación entre ellos. Un servicio de teleasistencia independientemente del servicio final ofrecido siempre parte (tal y como se mostró en la Figura 2.3) de una adquisición de datos del usuario que son evaluados y finalmente accesibles por los proveedores del servicio y por los cuidadores. Los proveedores proporcionarán el servicio de teleasistencia al usuario final y los cuidadores accederán a la información pudiendo intervenir directamente sobre la persona mayor si así lo desean. Los servicios que el proveedor proporciona pueden tener diferentes objetivos (ver Tabla 3.2) pero actuarán directamente sobre el usuario pretendiendo mejorar su calidad de vida.

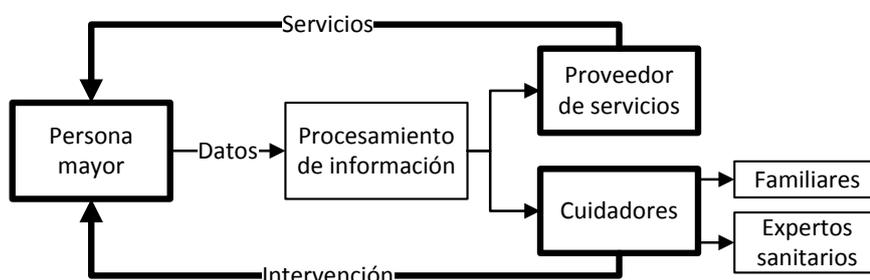


Figura 3.2. Agentes que intervienen en un servicio de teleasistencia

Existen en toda Europa, diferentes servicios de teleasistencia para personas mayores. La mayoría de ellos están basados en el uso de las TIC para hacer el estilo de vida más activo, saludable y social. Dentro del proyecto internacional SILHOUETTE (SILHOUETTE Partnership 2013b), cofinanciado por el programa INTERREG IVC y FEDER a través del mini-programa CREATOR, se identificaron y compartieron prácticas, conocimientos y herramientas con esta funcionalidad. El Apéndice C (sección C.1) muestra los principales resultados obtenidos durante la ejecución del proyecto. En todas las regiones se identificó un dispositivo similar al conocido en España como “botón del pánico”. Se trata de un brazalete/collar con un botón integrado que el usuario puede pulsar para generar alarmas. Esa alarma, dependiendo del caso, se

traduce en una llamada a un familiar o en un contacto con una centralita/servicio de emergencias. Este tipo de servicios de teleasistencia implican una respuesta por parte del usuario que debe pulsar el botón cuando necesita ayuda. Sin embargo, como se señala en (Ludwig et al. 2012), el usuario puede no estar capacitado para realizar la pulsación. Así, un tipo de servicio de teleasistencia que detecte situaciones de peligro de forma automática y sin intervención del usuario es mucho más interesante. Solamente la solución francesa “Geolocation” y la solución sueca “Check-up bag” supone una forma de monitorización del estilo de vida de la persona. Con la primera de las opciones se realiza una monitorización de la localización y con la segunda, una monitorización de las variables fisiológicas.

A nivel comercial, en el Apéndice C (sección C.2) se presentan dispositivos de monitorización de personas ampliamente extendidos y comercializados. Casi todas las soluciones encontradas (a excepción de “Comfort zone” y “Quiet care”) se focalizan en la monitorización de variables médicas y en el envío de esta información para su acceso y consulta por parte del propio usuario o de expertos médicos. Este tipo de soluciones basadas en el uso de las TIC facilitan tanto el trabajo como la comunicación con los expertos sanitarios (Pantelopoulos and Bourbakis 2010). A su vez, mejoran el control de síntomas, reducen costes y en definitiva, mejoran la calidad de vida de la persona monitorizada. Sin embargo, requieren de interacción por parte del usuario y que éste conozca el funcionamiento del sistema.

3.2.3. PROYECTOS DE INVESTIGACIÓN

Existen diferentes proyectos focalizados en la monitorización de personas mayores con diferentes objetivos. Entre todas las estrategias a nivel europeo creadas para facilitar la vida independiente de las personas mayores mediante el uso de las TIC, destaca el programa conjunto AAL (AAL JP 2015) tal y como se aprecia en la Tabla 3.3 adaptada de (Schreier and Amanatidou 2014).

El programa conjunto AAL se trata de un programa en colaboración entre la comisión europea y diferentes autoridades nacionales. Este programa (2008-2013) generó seis convocatorias de proyectos con el objetivo de desarrollar productos que facilitaran la vida independiente de las personas mayores desde un enfoque comercial. Todos los proyectos financiados fueron creados con el objetivo de poder ser comercializados en un plazo de 2 años tras la finalización del proyecto. Para ello dentro de los proyectos se debía garantizar la participación tanto de organismos de investigación como de asociaciones de usuarios y empresas interesadas en el producto final. En cada proyecto se deben involucrar al menos tres de los países elegibles en la convocatoria.

Tabla 3.3. Iniciativas europeas relacionadas con servicios de monitorización

Programa	Datos	Objetivo
Ambient Assisted Living Joint Programme (AAL JP) (AAL JP 2015)	Periodo: 2008-13 Proyectos: 17 Participantes: 130	Mejorar la calidad de vida y la autonomía de las personas mayores
Competitiveness and Innovation Programme, ICT Policy Support Programme (CIP-ICT PSP) (European Commission 2014)	Periodo: 2008-13 Proyectos: 12 Participantes: 144	Estimular la innovación y competitividad a través de la mejor del uso de las TIC por los ciudadanos, gobiernos y empresas
European Framework Programme FP7 (FP7-ICT) (European Commission 2013)	Periodo: 2008-15 Proyectos: 35 Participantes: 339	Diseño de sistemas de salud personalizados que mejoren la eficiencia en la asistencia
Joint Research Centre, Institute for Prospective Technological Studies (JRC) (European Commission 2015b)	Periodo: 2009-13 Proyectos: 3 Participantes: 3	Estudio y análisis de las necesidades, demandas y experiencias en sistemas de monitorización remota

Actualmente, dentro del programa Horizonte 2020 (European Commission 2015a) existen varias convocatorias centradas en la monitorización de personas mayores y en el diseño de nuevas soluciones para la salud y el bienestar. A su vez, el programa AAL ha iniciado una segunda convocatoria de proyectos buscando soluciones sostenibles (económicamente y socialmente) para mejorar la calidad de vida de las personas mayores promoviendo el autocuidado y la autonomía.

La sección del C.3 del Apéndice C recoge en el marco de las convocatorias descritas en la Tabla 3.3 los principales proyectos europeos en el ámbito de la monitorización y teleasistencia de personas mayores. La mayoría de estos proyectos se centran en una monitorización de variables médicas concretas (enfermedades cardiovasculares o respiratorias) y en una monitorización a nivel doméstico (dentro del hogar). Existen pocas soluciones basadas en una monitorización dentro y fuera del hogar. Entre este último tipo de soluciones destacaría el proyecto IS-ACTIVE y PAMAP. En ambos casos se realiza una monitorización del estilo de vida del usuario a partir de la medida y análisis de su comportamiento.

3.3. ANÁLISIS DEL ESTADO DEL ARTE

Los sistemas de monitorización de personas mayores basados en el uso de las TIC permiten la detección temprana, seguimiento y control de enfermedades/dolencias. Las dos tendencias principales en estos sistemas es la monitorización de variables médicas a través de sensores vestibles y el servicio de teleasistencia básico conocido como “botón del pánico”. En ambos casos, el servicio se basa en uno o varios dispositivos transmisores vestibles/portables y un dispositivo receptor que procesa la información recogida y decide la acción a ejecutar. El principal inconveniente es que el usuario debe conocer y ser capaz de interactuar con los sistemas de monitorización. En función de las características del usuario esto puede no ser factible. Este trabajo busca el diseño de un sistema de monitorización que pueda ser utilizado en cualquier usuario independientemente de sus capacidades tecnológicas y sus habilidades.

En general, los servicios de monitorización permiten incrementar la eficiencia y efectividad en la atención sanitaria a personas mayores suponiendo un decremento en los costes. Sin embargo, su baja penetración hace que sea necesario seguir trabajando en sistemas de monitorización de personas que permitan monitorizar su estilo de vida trabajando con usuarios y sobre periodos largos de tiempo.

Según (Ludwig et al. 2012) y (Heart and Kalderon 2013), los principales retos para los que los sistemas de monitorización de personas mayores penetren en la sociedad son:

- *Uso de tecnologías fiables.* Para la provisión de un servicio de teleasistencia es necesario el uso tanto de tecnologías como la implementación de una lógica fiable que permita ser utilizados en los procesos de cuidado del día a día de las personas mayores. Es necesario alcanzar un estado maduro de prototipo a través de un conjunto de pruebas que analice la tasa de fiabilidad del sistema.
- *Servicios adaptados a los requisitos de usuarios.* Los servicios ofrecidos deben adaptarse a los requisitos de caso de uso reales. Para ello es necesario mantenerse en contacto con los requisitos tanto de los usuarios finales como de todas las partes interesadas en el servicio (expertos técnicos, sanitarios, familiares...).
- *Integración con sistemas ya existentes.* Aunque la mayoría de los servicios de monitorización generan nuevas formas de visualización y de tratamiento de los datos, sería interesante conseguir su integración con los elementos ya existentes y conocidos por la comunidad. De esta forma se facilita su uso y su aceptación tanto por la comunidad científica como por los usuarios finales.

Los métodos propuestos en este trabajo parten del uso de tecnologías fiables y serán validados con varios usuarios durante periodos de tiempo prolongados. Cada uno de los métodos propuestos será validado tanto de forma independiente como de forma global. No se buscará el diseño de un nuevo dispositivo de monitorización sino el diseño de un método que pueda ser integrado en sistemas ya existentes con el objetivo de maximizar la aceptación.

El análisis DAFO (debilidades, amenazas, fortalezas y oportunidades) de la Figura 3.3, resume los resultados obtenidos en la revisión del estado del arte realizado en este capítulo. Esta figura ha sido construida a partir del análisis de todos los trabajos, soluciones comerciales y proyectos estudiados.

Debilidades	Amenazas
<ul style="list-style-type: none"> - Falta de integración en los sistemas ya existentes - Baja difusión entre la población - Necesidad de formación en la solución - Uso de tecnologías no fiables - Servicios no adaptados a los requisitos de los usuarios - Sistemas de monitorización válidos solo dentro del hogar 	<ul style="list-style-type: none"> - Soluciones centradas en trastornos específicos - Implantación de servicios no maduros - Necesidad de pruebas a gran escala - Reducida percepción de los beneficios por los usuarios
Fortalezas	Oportunidades
<ul style="list-style-type: none"> - Creciente penetración del uso de las TIC - Mejora en la calidad de vida - Reducción del coste y mejora de la eficiencia y efectividad de la atención sanitaria 	<ul style="list-style-type: none"> - Envejecimiento de la población - Buena disposición de los profesionales - Números proyectos de investigación en monitorización de comportamiento - Pocas soluciones comerciales

Figura 3.3. Análisis DAFO de los servicios de monitorización de personas

La principal debilidad de los sistemas de monitorización actuales reside en la no adaptación a los requisitos específicos de personas mayores y que hacen que los usuarios no sean conscientes de los beneficios que supone el uso del sistema (principal amenaza). Teniendo en cuenta la creciente penetración del uso de las TIC y los beneficios reales que acarrea su uso tanto en reducción de costes como en mejora de la calidad de vida del usuario (principales fortalezas), se pone de manifiesto la gran oportunidad que supone el uso de las TIC para la monitorización de usuarios.

Dentro de las principales tendencias existentes, cobran gran importancia los sistemas de monitorización basados en la medida del comportamiento ya que permiten la detección, el control y el seguimiento de varias enfermedades/dolencias mediante el análisis de anomalías de comportamiento. La utilidad de estos sistemas parte de la profunda relación que existe entre el estado de salud del usuario y su comportamiento. Tal y como puede verse en la Figura 3.4, adaptada de (Kivel 2013), el comportamiento es uno de los principales elementos que definen el estado de salud de una persona. El comportamiento ha de entenderse como el conjunto de actividades que el usuario realiza en su día a día y a través de las cuales se define su estilo de vida.

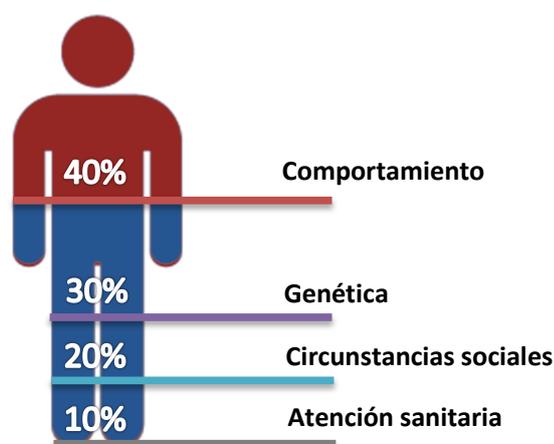


Figura 3.4. Elementos que definen el estado de salud de una persona

El diseño de un sistema de monitorización que no requiera de interacción por parte del usuario y que se base en el análisis del comportamiento ha sido abordado por un número muy pequeño de trabajos y proyectos de investigación. La mayoría de ellos, al igual que los servicios comerciales de teleasistencia, se basan principalmente en la monitorización de variables de salud.

Una de las principales limitaciones encontradas en el estado del arte es la necesidad de una etapa de experimentación que permita validar los sistemas propuestos. En un 80% de los trabajos analizados en (Orwat, Graefe, and Faulwasser 2008) no existía una etapa de experimentación y validación documentada.

Los métodos de medida, modelado y detección propuestos en este trabajo serán utilizados para detectar y controlar cambios de comportamiento. Se diseñará una etapa de experimentación en entorno real que permita realizar su validación y evaluar su aplicabilidad para la monitorización de enfermedades como demencia, Alzheimer, insomnio y/o el síndrome de fragilidad.

3.4. CONCLUSIONES

En este capítulo se presenta el estado del arte de las principales tendencias existentes en la monitorización de personas mayores. La monitorización de variables médicas y el servicio conocido como “botón del pánico” son las principales alternativas. Sin embargo, cada vez existe un mayor número de trabajos centrados en la monitorización del comportamiento.

La principal debilidad de los sistemas de monitorización actuales es su baja aceptación por las personas mayores que lo conciben como un elemento intruso en su día a día. Esto hace que los usuarios no conciban las ventajas que un sistema de monitorización puede introducir en su vida y sean reacios a su uso y aceptación. Sin embargo, tal y como se ha puesto de manifiesto en diferentes estudios existentes en el estado del arte, un sistema de monitorización basado en TIC permite reducir costes en atención sanitaria y mejorar simultáneamente la calidad de vida. Así, este tipo de sistemas presentan una serie de fortalezas y oportunidades que deberán ser aprovechadas para combatir sus debilidades y amenazas.

En este trabajo se propone el diseño de un sistema de monitorización basado en la detección de anomalías de comportamiento y que servirá para la detección, el control y el seguimiento de problemas relacionados con trastornos de comportamiento. El objetivo del sistema será mejorar la calidad de vida de los usuarios. Este sistema supone un avance respecto a las soluciones actuales que se centran principalmente en la monitorización de variables médicas y en el servicio “botón del pánico”.

El sistema propuesto en este trabajo se guía por los requisitos necesarios para garantizar la aceptación del usuario, no seguidos en la mayoría de las soluciones existentes y se compone de tres métodos principales: método de medida del comportamiento, método de modelado del comportamiento y método de detección de anomalías.

CAPÍTULO 4.

MÉTODO DE MEDIDA: ACTIVIDAD FÍSICA

4.1. INTRODUCCIÓN

Un método de medida del comportamiento humano permite monitorizar las actividades de la vida diaria del usuario siendo ésta una de las principales características que definen el estado de salud de una persona. Esta medida cobra gran importancia en la monitorización de personas mayores ya que permite garantizar su seguridad y mejorar su calidad de vida (Gokalp and Clarke 2013).

El objetivo de este capítulo se describe gráficamente en la Figura 4.1 y es el “Diseño y validación de un método de medida”. Siguiendo el esquema de la Figura 2.4, el método propuesto debe permitir la medida del comportamiento mediante el diseño de un método de procesado (RF-2) que a partir de la información adquirida de un sensor (RF-1), obtenga una magnitud válida de medida (RF-3).

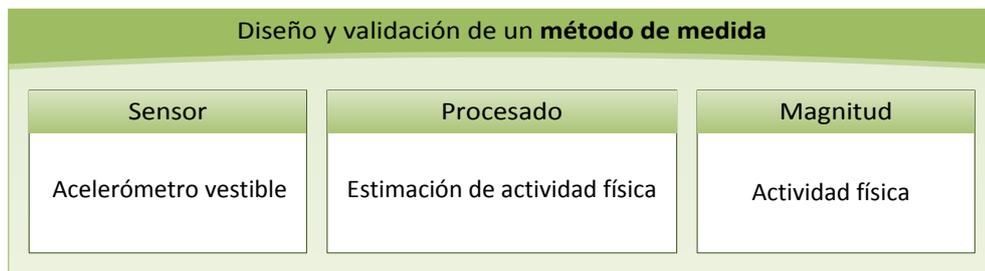


Figura 4.1. Objetivo principal del Capítulo 4

Para conseguir el objetivo anteriormente descrito, en este capítulo se estudiará el estado del arte en sensores para la medida del comportamiento, en métodos de procesado y en magnitudes de medida.

En la sección 4.2 se presentan los principales sensores utilizados en el estado del arte para la medida del comportamiento. Dentro de este tipo de sensores destacan los sensores de presencia, movimiento, localización y/o videovigilancia. Estos sensores pueden ser clasificados en dos categorías principales: sensores ambientales y sensores vestibles. Los primeros son menos invasivos y más sencillos de utilizar mientras que los segundos son más precisos pero presentan una alta tasa de rechazo por parte de los usuarios. En este trabajo se propone el uso de un sensor de movimiento vestible integrado en un reloj junto con un método de adquisición pasivo (sin interacción por parte del usuario). De esta forma se consigue un compromiso entre invasividad y precisión.

Las principales magnitudes utilizadas en el estado del arte para la medida del comportamiento son también presentadas en la sección 4.2. La localización, medida de presencia, consumo eléctrico, uso de elementos y actividad física son las principales magnitudes encontradas en el estado del arte para la medida del comportamiento. En este trabajo se opta por el uso de la actividad física como magnitud para la medida. Se trata de una magnitud estándar, validada y ampliamente utilizada por la comunidad científica con diversos objetivos de monitorización. Sin embargo, su uso para la monitorización de personas mayores es escaso y destaca principalmente con fines deportivos en adultos.

De entre todos los métodos disponibles para la medida de la actividad física y que son descritos en detalle en la sección 4.3 (métodos estándar, métodos objetivos y métodos subjetivos) se opta por el uso de acelerómetros vestibles. Este tipo de sensores miden la aceleración del cuerpo que los porta, requieren una etapa compleja de procesado para su transformación a medida de actividad física y generalmente tienen una autonomía corta. En este trabajo se propone un nuevo método de estimación de la actividad basado en el uso de la aceleración y que permite reducir la complejidad del procesado y en consecuencia el consumo eléctrico del mismo (tiempo de vida).

La sección 4.4 presenta el diseño y validación del método de medida propuesto en este trabajo. En el diseño de este método se tienen en cuenta tanto los requisitos funcionales como los no funcionales (ver Capítulo 2). Pese a que el diseño tiene lugar siguiendo requisitos específicos para su uso en personas mayores, la validación técnica del método de medida tiene lugar con usuarios adultos para realizar una validación del método con alta variabilidad de actividad.

4.2. ESTADO DEL ARTE:

MEDIDA DEL COMPORTAMIENTO

A continuación se muestra un conjunto de tablas (de Tabla 4.1 a Tabla 4.8) mostrando las principales tendencias existentes en el estado del arte para la medida del comportamiento humano: sensores PIR de movimiento, sensores de localización, cámaras, detectores de consumo eléctrico, etiquetas RFID, detectores magnéticos, sensores de presión y medidores de cantidad de movimiento. En cada tabla se describe el tipo de sensor, la magnitud usada para medir el comportamiento y varias referencias de su uso en el estado del arte. Según (S Brownsell and Bradley 2011), la tendencia más utilizada es el uso de sensores PIR detectores de movimiento (descritos en la Tabla 4.1).

Tabla 4.1. Medida del comportamiento: sensores PIR

Sensor	Tipo	Magnitud
Sensores PIR detectores de movimiento	Ambiental	Presencia
	Descripción	
	Sensores piro eléctricos que a través de la medida del movimiento detectan presencia. Integrados en el hogar permiten localizar usuarios e identificar su presencia en cada dependencia	
	Referencias	
(Fouquet and Franco 2010) (Fleury et al. 2010) (Virone 2009)		

Tabla 4.2. Medida del comportamiento: sensores de localización

Sensor	Tipo	Magnitud
Sensores de localización inalámbrica	Ambiental o vestible	Localización
	Descripción	
	Sensores que porta el usuario y que permiten realizar su localización. Implica el despliegue de una red de sensores en el hogar junto con la necesidad de que el usuario porte un sensor en todo momento	
	Referencias	
(Marco et al. 2008) (Charlon, Fourty, and Campo 2013)		

Tabla 4.3. Medida del comportamiento: vídeo y/o micrófonos

Sensor	Tipo	Magnitud
Cámaras de vídeo y/o micrófonos	Ambiental	Vídeo, sonido
Descripción		
Sensores de vídeo y/o sonido configurados para constituir un sistema de videovigilancia que permita realizar una monitorización de las actividades del usuario		
Referencias		
(Fleury et al. 2010) (W. A. Khan et al. 2012)		

Tabla 4.4. Medida del comportamiento: detectores de consumo eléctrico

Sensor	Tipo	Magnitud
Detectores de consumo eléctrico	Ambiental	Consumo eléctrico
Descripción		
Sensores integrados en dispositivos eléctricos que permiten monitorizar el consumo de los usuarios y a través de ellos monitorizar su comportamiento		
Referencias		
(Fleury et al. 2010) (W. A. Khan et al. 2012)		

Tabla 4.5. Medida del comportamiento: etiquetas RFID

Sensor	Tipo	Magnitud
Etiquetas RFID	Ambiental	Uso de elementos
Descripción		
RFID es un sistema de almacenamiento y recuperación de datos que es utilizado para monitorizar el comportamiento del usuario a través de la identificación de los objetos que utiliza		
Referencias		
(Kasteren et al. 2010) (Fleury et al. 2010) (Chen, Cook, and Crandall 2013) (Chen et al. 2013)		

Tabla 4.6. Medida del comportamiento: detectores magnéticos

Sensor	Tipo	Magnitud
Detectores magnéticos	Ambiental	Uso de elementos
	Descripción	
	Detección de cierre y apertura de elementos mecánicos. Permite monitorizar el uso de elementos como ventanas, puertas, frigorífico, horno...	
	Referencias	
(Kasteren et al. 2010) (Fleury et al. 2010) (Wang et al. 2011)		

Tabla 4.7. Medida del comportamiento: sensores de presión

Sensor	Tipo	Magnitud
Sensores de presión	Ambiental	Uso de elementos
	Descripción	
	Sensores que pueden ser integrados en objetos cotidianos como alfombras, camas, sillas... para detectar el uso que el usuario realiza del objeto y a través de él, monitorizar su comportamiento	
	Referencias	
(Fleury et al. 2010) (Virone 2009)		

Tabla 4.8. Medida del comportamiento: medidores de cantidad de movimiento

Sensor	Tipo	Magnitud
Medidores de cantidad de movimiento	Vestible	Actividad física
	Descripción	
	Principalmente acelerómetros y podómetros que el usuario porta y que permiten medir la actividad física que realiza	
	Referencias	
(Fontecha and Hervás 2011) (Hajj Chegade et al. 2012) (Fleury et al. 2010) (Doukas and Maglogiannis 2008)		

Todos los sensores mostrados en las tablas anteriores pueden ser englobados en dos categorías principales (ya mencionadas en la sección 2.3.2): sensores ambientales y sensores vestibles. El uso de cada uno de estos tipos de sensores conlleva ventajas e inconvenientes que han de ser tenidos en cuenta cuando se diseña un sistema para la monitorización de personas mayores.

De acuerdo a toda la información recopilada, en la Tabla 4.9 se presenta un resumen de las dos principales alternativas (sensores ambientales y sensores vestibles) diferenciando en cada categoría los sensores presentados anteriormente. En la revisión del estado del arte realizada en (Orwat et al. 2008), prácticamente la mitad de los trabajos analizados se decantó por cada una de las dos alternativas. No existe una tendencia y predilección clara sobre qué tipo de sensores son mejores para la monitorización de personas.

Tabla 4.9. Tipos principales de sensores para la adquisición de medidas

Tipo	Sensores ambientales	Sensores vestibles
Definición	Sensores integrados en objetos o localizaciones comunes para el usuario formando sistemas de computación pervasiva	Sensores que pueden ser portados por el usuario mediante su integración en elementos vestibles
Implementación	Integración en el entorno creando espacios inteligentes	Integrados de forma transparente en prendas o elementos vestibles
Ventajas	<ul style="list-style-type: none"> - Poco invasivos - Información multimodal 	<ul style="list-style-type: none"> - Precisión - Válidos en entornos multiusuario
Inconvenientes	<ul style="list-style-type: none"> - No válidos en entornos multiusuario - Complejidad de procesamiento - Vulneración de la privacidad - Instalación y mantenimiento complejo 	<ul style="list-style-type: none"> - Reluctancia a portarlos - Invasivos
Ejemplos	<ul style="list-style-type: none"> - Detectores de movimiento - Cámaras de vídeo y/o micrófonos - Detectores de consumo eléctrico - Detectores magnéticos - Sensores de presión 	<ul style="list-style-type: none"> - Sensores de localización inalámbrica - Etiquetas RFID - Medidores de cantidad de movimiento

Los sensores ambientales permiten recolectar información multimodal y son poco invasivos para el usuario al encontrarse integrados en el entorno (Chikhaoui, Wang, and Pigot 2012) (Barnes, Webster, and Mizutani 2006). Sin embargo no son válidos en entornos multiusuario al ser incapaces de diferenciar fácilmente el comportamiento entre usuarios (Gokalp and Clarke 2013). En algunos casos, como en el de sensores de visión, además de la intromisión y las connotaciones que para el usuario conlleva su monitorización, debe tenerse en cuenta la complejidad en el tratamiento de los datos (Gokalp and Clarke 2013). La privacidad del usuario cobra una mayor importancia para los sensores ambientales cuando se considera la posibilidad de monitorizar el comportamiento del usuario en el dormitorio o en el baño.

Los sensores vestibles son más precisos que los sensores ambientales pese al ruido que en ocasiones introduce el movimiento del cuerpo humano (Gokalp and Clarke 2013) y la información recolectada puede ser utilizada tanto para predecir como para prevenir riesgos de salud (Chan et al. 2009). En muchas ocasiones resultan invasivos para el usuario (Chikhaoui et al. 2012) produciéndose una reluctancia a portarlos (Dias et al. 2012). Este hecho se agrava en el caso de usuarios con problemas cognitivos para los cuales existe el temor a que el usuario no recuerde colocar el sensor y en determinados casos de que lo pierda o se desprenda de él. En función de la naturaleza del sensor, también debe tenerse en cuenta la imposibilidad de vestirlo en determinadas circunstancias y durante la realización de algunas actividades (ducha, cama, nadar...). Los sensores vestibles han sido validados para la medida del comportamiento mediante el reconocimiento de actividades (caminar, correr, reposo, caídas, inactividad...) y la medida de la actividad física (Wang et al. 2011) (Heil et al. 2012).

Dentro de los sensores vestibles destacan principalmente las cuatro alternativas mostradas en la Tabla 4.10 (S. Khan et al. 2012), (Pantelopoulos and Bourbakis 2010): sensores integrados en complementos, en prendas de vestir, teléfonos móviles y dispositivos específicos.

Los teléfonos móviles son considerados en algunos trabajos como elementos vestibles ya que suelen ser portados por los usuarios casi todo el día (Joe and Demiris 2013). Sin embargo, su uso tiene varias limitaciones debido al tipo de sensores que suelen tener integrados (baja precisión y baja tasa de muestreo disponible) y a la variabilidad en su posición dependiendo del usuario (Abbate et al. 2012). Otro de los inconvenientes en el uso de teléfonos móviles como dispositivos vestibles radica en que se trata de dispositivos generalmente pesados y de mayor tamaño que otro dispositivo vestible que solamente integrara los sensores requeridos para la medida.

Tabla 4.10. Principales alternativas de sensores vestibles

Sensor	Descripción
Integrado en complementos	Los sensores se integran en complementos comúnmente vestidos por el usuario. Destacan principalmente collares, cinturones y relojes
Integrado en prendas de vestir	De forma transparente para el usuario la adquisición de variables se realiza mediante las prendas que éste porta. Destaca el uso de camisetas y sensores integrados en calzado
Teléfono móvil	En algunas referencias, el teléfono móvil es considerado como un ejemplo de sensor vestible. La adquisición de variables se realiza a través de los sensores que éste lleva integrados
Dispositivos específicos	Dispositivos, principalmente relacionados con la salud, como tensiómetros o pulsímetros también son considerados como ejemplos de sensores vestibles

En el uso de dispositivos específicos para la monitorización destaca sobre todo el uso de sensores de salud. La principal ventaja de este tipo de dispositivos en su eficiencia y fiabilidad ya que se trata, en general, de dispositivos ampliamente utilizados y conocidos por la comunidad científica. Sin embargo, en muchos casos su portabilidad no está clara ya que son ampliamente invasivos y no pueden ser utilizados para la monitorización constante como es el caso del uso de tensiómetros.

Existen varias soluciones comerciales de dispositivos específicos que pueden ser utilizados para la medida del comportamiento humano. El Apéndice C (sección C.4) muestra parte de los sensores que se encuentran actualmente en el mercado. Como se puede comprobar, la mayor parte de estos sensores se integran en dispositivos con forma de pulsera y con propósitos deportivos. Ninguno de ellos ha sido especialmente diseñado para la monitorización del comportamiento de personas mayores teniendo en cuenta las características especiales de este grupo de usuarios.

Actualmente, la penetración de estos dispositivos en el mercado es muy alta. Según el estudio realizado por Endeavour Partners en Estados Unidos (Ledger 2014), un 10% de las personas mayores de 18 años son actualmente consumidores de alguno de los dispositivos mencionados. El 25% de las personas que poseen este tipo de dispositivos según el mismo estudio (Ledger 2014) se encuentran dentro del rango de 25 a 34 años y su uso es principalmente para la realización y monitorización de actividades

deportivas. Un pequeño porcentaje de las personas (7%) se encuentra dentro del rango de 55 a 64 años siendo en este caso el objetivo de mejorar su estado de salud y calidad de vida.

La solución a los principales inconvenientes que plantean los sensores vestibles (invasividad y reluctancia de los usuarios) pueden resolverse mediante la integración de los sensores en prendas de vestir o complementos que sean comúnmente vestidos por el usuario (Steele et al. 2009). De esta forma se consigue una medida fiable (RT-1) de forma transparente y sin necesidad de disminuir su confort (RU-1, RU-4). Esta solución también resuelve los problemas descritos en relación a los usuarios con problemas cognitivos, se reduce la intromisión en la privacidad del usuario y se consiguen medidas más precisas que en el caso de sensores ambientales. También ha de tenerse en cuenta la disminución de la complejidad en la instalación y el mantenimiento de los sensores vestibles frente a los sensores ambientales. Esta determinación fue una de las conclusiones extraídas en la ejecución del proyecto ATLAS dirigido por Fundación CTIC (Fundación CTIC 2010).

De acuerdo a estas ideas, en este trabajo se propone el uso de un dispositivo vestible con forma de reloj debido a las reticencias que se describen en algunas referencias respecto al uso de sensores integrados en prendas de vestir (Schaar and Ziefle 2011). Este tipo de dispositivos para la medida de la actividad física han sido validados en varios trabajos como (Holzinger and Searle 2010) o (Zhang et al. 2012).

Según un estudio realizado en 2010 en Reino Unido (Mintel 2010) alrededor de un 86% de la población posee y utiliza un reloj. Este porcentaje se incrementa en las personas mayores que suelen llevarlo puesto durante la mayor parte del día. Así, su uso como dispositivo para la monitorización permite reducir la invasividad (RU-1, RU-2 y RU-3) reduciendo la reluctancia del usuario a ser monitorizado.

De entre todas las magnitudes presentadas en las tablas anteriores (de Tabla 4.1 a Tabla 4.8), la actividad física es la única magnitud estándar y fácilmente interpretable. Pese a esto, su uso para medir el comportamiento está poco extendido en personas mayores y se focaliza principalmente con fines lúdicos y deportivos en los sensores comerciales presentados en el Apéndice C (sección C.4).

El estudio del estado del arte en sensores y magnitudes para la medida del comportamiento pone de manifiesto la utilidad de sensores integrados en complementos vestibles y las ventajas del uso de una magnitud estándar como la actividad física.

4.3. ACTIVIDAD FÍSICA

4.3.1. DEFINICIÓN E IMPORTANCIA DE LA ACTIVIDAD FÍSICA

La actividad física (*physical activity*, PA) se define como cualquier movimiento del cuerpo humano que se traduce en un gasto energético (Caspersen et al. 1985). La medida de la actividad física es una importante medida cuantitativa que permite detectar desórdenes de comportamiento como hiperactividad y/o insomnio. También, a través de la medida de la actividad física se pueden detectar comportamientos sedentarios que podrían contribuir al desarrollo de importantes patologías como ataques cardiovasculares o artrosis. Así, además de ser una magnitud conocida por la comunidad científica y utilizada en numerosos trabajos científicos, la actividad física es considerada una de las tres claves más importantes en el envejecimiento activo de la población (Oxley 2009) y una de las herramientas más eficaces para la reducción de la mortalidad y la prevención de enfermedades crónicas (King and King 2010).

La Figura 4.2, adaptada de (Vanhees et al. 2005), muestra la relación existente entre actividad física y salud. A través de las características específicas que definen a un usuario, existe una relación de doble vía entre actividad física y salud. Por un lado, la actividad física puede ser concebida como una herramienta para la prevención de enfermedades (Vogel et al. 2009) y por otro lado, a través de la actividad física puede realizarse la detección de patologías relacionadas con desórdenes de actividad como el insomnio (Weiss et al. 2010) o la hiperactividad (Bussmann, Ebner-Priemer, and Fahrenberg 2009).

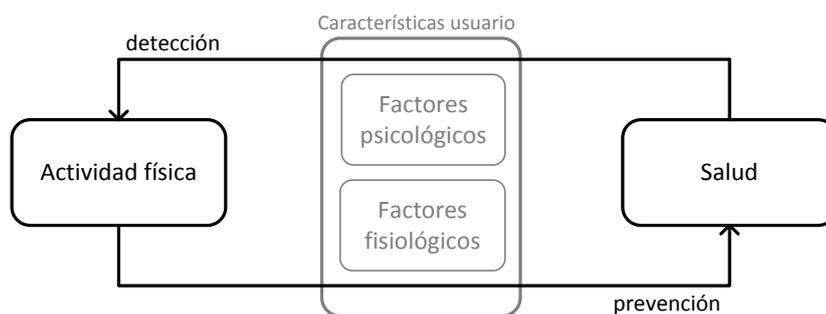


Figura 4.2. Relación entre actividad física, condiciones físicas y salud

Las características que definen a un usuario pueden dividirse en factores psicológicos y en factores fisiológicos. Los factores psicológicos resumen las características psicológicas de los usuarios y están relacionados con su edad y su desarrollo cognitivo. Los factores fisiológicos hacen referencia a características como el peso, la estatura y/o el índice de masa corporal. Todos estos factores definen a un usuario y son

condicionantes en la relación bidireccional existente entre actividad física y salud. Un usuario obeso podrá mejorar su salud a través de la realización de actividad física previniendo enfermedades cardiovasculares y patologías directamente ligadas con la obesidad. Así, a través de la actividad física el usuario prevendrá enfermedades pero también modificará sus características fisiológicas. Por otro lado, la monitorización de la actividad física en ese mismo usuario obeso será de utilidad para detectar un comportamiento sedentario y motivar, por ejemplo, la realización de actividades deportivas. A través de la monitorización de la actividad física se detecta un posible problema de salud y a través de su realización se puede prevenir el mismo.

En (Berlin, Storti, and Brach 2006) se indican las principales utilidades de la medida de la actividad física y se pone de manifiesto su importancia con dos objetivos principales: monitorización y detección. La actividad física sirve para la monitorización de planes de ejercicios y de recomendaciones evaluando la adherencia del usuario a indicaciones dadas por expertos. Sin embargo, la medida y evaluación de la actividad física permite detectar situaciones críticas como periodos de inactividad y sedentarismo.

La definición dada de actividad física (“cualquier movimiento del cuerpo humano que se traduce en un gasto energético”) no debe confundirse con los términos “ejercicio” o “deporte” (Caspersen et al. 1985). Estos últimos términos deben interpretarse como movimientos planeados, estructurados y repetitivos del cuerpo humano con el objetivo de mejorar o mantener la forma física.

4.3.2. MÉTODOS PARA LA MEDIDA DE LA ACTIVIDAD FÍSICA

En la literatura científica, existen diversos métodos para realizar la medida de la actividad física (Plasqui, Bonomi, and Westerterp 2013). Estos métodos pueden ser divididos en tres clases diferenciadas cuyas ventajas, inconvenientes y aplicaciones se presentan a continuación:

- *Métodos estándar (ver Tabla 4.11)*. Son los más exactos pero los más complicados de utilizar en un entorno real. Destaca la técnica del agua doblemente marcada (*doubly labelled water, DLW*) considerado como el *gold-standard* de la medida.
- *Métodos objetivos (ver Tabla 4.12)*. Destaca el uso de sensores de movimiento. Se trata de una de las técnicas más validadas y extendidas en la medida de la actividad física.
- *Métodos subjetivos (ver Tabla 4.13)*. Engloban principalmente el uso de cuestionarios. Es una técnica de bajo coste fácilmente aplicada en estudios con poblaciones de gran tamaño.

Tabla 4.11. Métodos estándar de medida de actividad física

Tipo	Método	Ventajas	Inconvenientes
Métodos estándar	Observación directa (Ainsworth et al. 2000)	- Información cualitativa y cuantitativa - Información contextual	- Costoso en tiempo y esfuerzo - Usuarios reacios a ser observados
	Técnicas calorimétricas indirectas (Haugen, Chan, and Li 2007)	- Exactitud y precisión	- Invasivo - Alto coste - Válido únicamente para estudios en laboratorio
	Agua doblemente marcada (Schoeller 1988)	- Exactitud y precisión - No invasivo y transparente para el usuario	- Invasivo - Alto coste - Imposibilidad de detectar patrones, solo se mide la actividad total

Tabla 4.12. Métodos objetivos de medida de actividad física

Tipo	Método	Ventajas	Inconvenientes
Métodos objetivos	Sensores fisiológicos (Hautala et al. 2010)	- Buena relación con la actividad física - Se puede detectar intensidad, frecuencia y duración de la actividad - Recogida de datos sencilla	- Útil principalmente para actividades aeróbicas - Calibración dependiente del individuo
	Sensores de movimiento (Welk 2002)	- Proporciona información de intensidad, frecuencia y duración de la actividad - Bajo coste - No invasivo - Recogida de datos fácil, rápida y durante periodos largos	- Baja exactitud en las medidas - Inexistencia de ecuaciones exactas de relación con la actividad física - Limitaciones para estudios grandes (tiempo de vida limitado)

Tabla 4.13. Métodos subjetivos de medida de actividad física

Tipo	Método	Ventajas	Inconvenientes
Métodos subjetivos	Cuestionarios	<ul style="list-style-type: none"> - Se puede detectar intensidad, frecuencia y duración de la actividad - Bajo coste 	<ul style="list-style-type: none"> - Problemas con la validez, fiabilidad y repetición de los experimentos - Problemas con la interpretación de la actividad física por parte de los usuarios - Poca adherencia y aceptación de usuarios

De acuerdo al resumen realizado en las tablas anteriores (Tabla 4.11, Tabla 4.12 y Tabla 4.13), la Figura 4.3 muestra los principales métodos representados en una gráfica según su facilidad de uso y la precisión de sus medidas. Hay dos métodos que destacan principalmente: la técnica del agua doblemente marcada (*gold-standard*) y los sensores de movimiento. En el primer caso destaca la precisión del método y en el segundo su fácil uso y aceptación por parte de los usuarios.

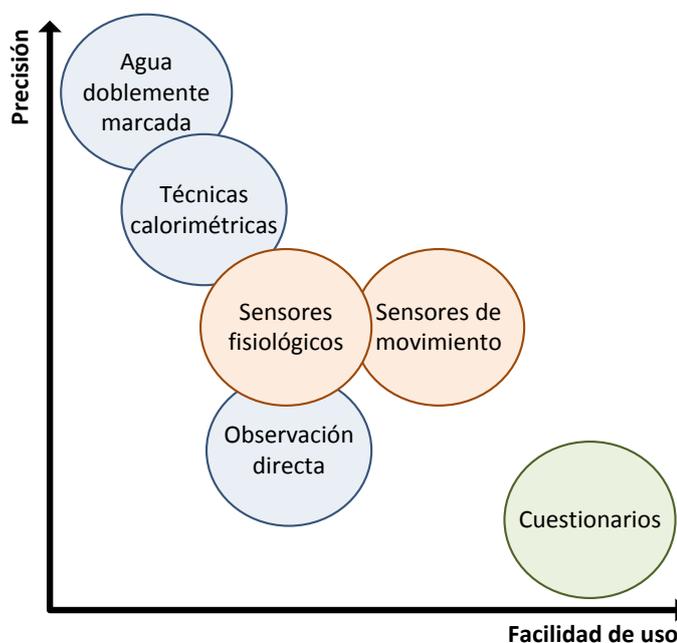


Figura 4.3. Características de los principales métodos para medir la actividad física

La técnica del agua doblemente marcada implica un alto coste y aunque su medida es transparente para el usuario, requiere de análisis biomédicos de forma periódica y se convierte en una medida que requiere un alto esfuerzo. Por otro lado, este tipo de técnicas permiten detectar el total de calorías consumidas pero no permiten detectar patrones de actividad física en periodos cortos de tiempo.

Dentro de los sensores de movimiento para medir la actividad física, destacan los podómetros y los acelerómetros:

- *Podómetros*: Un podómetro es un dispositivo electrónico o electromecánico, ligero, de pequeño tamaño y portátil. Generalmente, deben ser colocados en la cintura y cerca de la cadera. Su funcionamiento se basa en la detección y cuenta cada paso que una persona da a través del movimiento de sus caderas. A través de la cuenta de pasos, se realiza la medida de la actividad física. La principal limitación de este tipo de dispositivos reside en la realización de actividad sin movimiento vertical (que es precisamente el que miden los podómetros) y en la imposibilidad de proporcionar información acerca de la intensidad del movimiento. No resultan adecuados para la monitorización de determinadas actividades como puede ser el transporte de cargas o deporte como ciclismo, natación...
- *Acelerómetros*: Su uso como medidor de actividad física se realiza a través de la relación existente entre los movimientos que realiza una persona (medida a través de la aceleración) y las fuerzas musculares que los ocasionan. La relación entre estas fuerzas musculares y la actividad física reside en la propia definición de esta magnitud. La existencia de acelerómetros de varios ejes (vertical, medio-lateral y anterior-posterior) permite realizar una mejor caracterización del movimiento que el realizado por los podómetros. Están generalmente formado por elementos piezoeléctricos y un microcontrolador encargado de cuantificar la aceleración en cada uno de los ejes de medida del acelerómetro a través de la magnitud adimensional *counts*.

Actualmente, el uso de sensores de movimiento y en particular, el uso de acelerómetros se encuentra ampliamente extendido y es fácil encontrarlos formando parte de diferentes dispositivos electrónicos como pueden ser teléfonos móviles, cámaras de fotos u ordenadores. Un acelerómetro es un dispositivo capaz de medir la aceleración que experimenta un cuerpo. Este tipo de aceleración no tiene por qué ser debida a la aceleración de coordenadas (cambio de velocidad de un cuerpo en el espacio), sino que es el tipo de aceleración asociada con el fenómeno de peso. Así, un

cuerpo en reposo posee una aceleración de $9,81 \text{ m/s}^2$ (aceleración en caída libre) que se utiliza como referencia y se nombra con la letra g .

El uso de acelerómetros para la medida de la actividad física también es muy común sobre todo por tratarse de un dispositivo poco invasivo y de bajo coste cuya información puede ser recolectada de forma sencilla y cuya fiabilidad como medidor de actividad física está demostrado (Murphy 2009).

La principal limitación en el uso de acelerómetros para la medida de la actividad física está relacionada con su tiempo de vida según varios trabajos como (Ward et al. 2005) o (Chen and Bassett 2005). Para que un dispositivo pueda ser utilizado en la monitorización de personas mayores mediante la medida de la actividad física es necesario según (Plasqui et al. 2013) que el tiempo de vida del dispositivo sea de alrededor de 3 días. Para conseguir tiempos de vida grandes, es necesario aumentar el tamaño de la batería aumentando así el tamaño del sensor. En (Avci and Bosch 2010) y (Plasqui et al. 2013) se señala el sistema de alimentación como uno de los principales limitadores en la miniaturización de los dispositivos de medida de actividad física y por tanto uno de los principales motivos de la invasividad de la medida. Otro de los principales inconvenientes en el uso de acelerómetros es su capacidad de memoria (Avci and Bosch 2010) (Ward et al. 2005) (Chen and Bassett 2005). Es necesario que el dispositivo que integra el acelerómetro tenga memoria suficiente para el almacenamiento de datos durante al menos 2 días.

4.4. DISEÑO DEL MÉTODO DE MEDIDA

4.4.1. SELECCIÓN DE UNA MAGNITUD (RF-3)

Del estudio del estado del arte realizado en la sección 4.2 se determinó a la actividad física como la magnitud más fiable para la medida del comportamiento. Este hecho parte de que la actividad física es una magnitud estándar y ampliamente utilizada como medida del comportamiento sobre todo en el ámbito deportivo.

En este trabajo se parte del uso de la actividad física como medida de comportamiento con el objetivo de realizar una monitorización global del usuario. La monitorización de la actividad física permitirá caracterizar adecuadamente la actividad realizada por las personas mayores.

4.4.2. SENSOR DE MEDIDA (RF-1)

En el caso de personas mayores, se suele optar por el uso de sensores de tipo ambiental (generalmente sensores de presencia) para realizar una monitorización de su comportamiento. En estos casos, su comportamiento se mide a través de las habitaciones por las que el usuario se mueve y los elementos del hogar con los que interactúa. Este tipo de sistemas no permiten una monitorización continua tanto dentro como fuera del hogar y su uso se restringe a la detección de situaciones muy específicas. Un sensor de tipo vestible integrado en complementos, tal y como se describió en la sección 4.2, conseguiría resolver esta limitación y permitir una monitorización del usuario dentro y fuera del hogar.

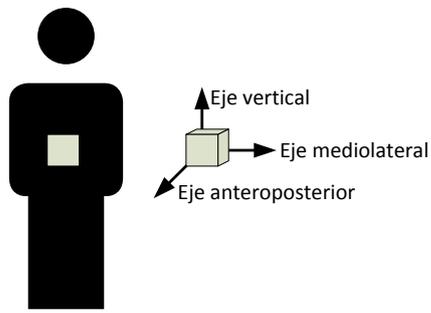
De entre todos los métodos de medida de la actividad física propuestos en la sección 4.3.2, destaca el uso de sensores de movimiento al permitir realizar una medida objetiva de la actividad caracterizando tanto la duración como la intensidad y la frecuencia. A su vez son sensores poco invasivos, vestibles y que permiten monitorizar periodos de tiempo prolongados.

El uso de acelerómetros como medidor de actividad física es claramente destacable sobre el uso de podómetros. Se trata de una de las técnicas más antiguas y más extendidas en la medida de la actividad física y cumple los requisitos descritos en la sección 2.5 (monitorización no invasiva y autónoma) para su uso por personas mayores. Existen estudios que validan su uso y que prueban la existencia de una relación lineal entre su medida y la energía consumida (Bouten et al. 1997) (Welk 2002). Las principales limitaciones de los acelerómetros en su uso están relacionadas con su consumo eléctrico y con la inexistencia de ecuaciones estándar que permitan transformar de forma universal la aceleración en actividad física.

Los acelerómetros además de poder ser utilizados para la medida de la actividad física son señalados en (Koeneman et al. 2011) como uno de los sensores más adecuados para la monitorización de personas mayores de acuerdo a su precisión y fiabilidad. Sin embargo, generalmente son utilizados para la identificación de actividades y no para la medida de la actividad física como se propone en este trabajo.

La Tabla 4.14 muestra el resumen de las características que debería de cumplir un acelerómetro para poder medir la actividad física según varios trabajos como (Bouten et al. 1997) o (Mathie et al. 2004). Estos requisitos son seguidos por todos los trabajos basados en la medida de actividad mediante el uso de acelerómetros.

Tabla 4.14. Características de un acelerómetro para medir actividad física

Característica	Valor	Descripción
Frecuencia de adquisición	> 40 Hz	La mayoría de los movimientos del cuerpo humano se encuentran en el rango de 0,3 a 3,5 Hz pero para conseguir una caracterización completa del movimiento se necesita medir el rango hasta 20Hz (Mathie et al. 2004). Siguiendo el teorema de Nyquist (Oppenheim, Willsky, and Nawab 1983) es necesario una frecuencia de adquisición de medida superior a 40Hz.
Rango de medida	> ±6 g	Un acelerómetro con un rango de medida entre -6g y +6g consigue realizar una caracterización completa del movimiento (Bouten et al. 1997) (Chen and Bassett 2005).
Número de ejes	3	 <p>El acelerómetro debe ser triaxial midiendo la aceleración en los 3 ejes del movimiento: vertical, mediolateral y anteriorposterior.</p>

En (Atallah et al. 2010) se realiza un análisis sobre las posibles colocaciones de un acelerómetro sobre el cuerpo humano (oreja, pecho, muñeca, rodilla, tobillo, brazo y cintura) y su validez para la medida de la actividad física en la realización de actividades de la vida diaria (ADL). En este análisis se realiza el estudio en la clasificación del nivel de actividad en función de la localización de los sensores:

- *Intensidad muy baja (reposo)*. La tasa de precisión es razonable para todas las localizaciones propuestas pero la sensibilidad es baja para todas las localizaciones. La muñeca y la oreja proporcionan los mejores resultados.
- *Intensidad baja*. El sensor colocado en la cintura proporciona la máxima precisión y sensibilidad de todo el estudio realizado.
- *Intensidad media*. El sensor de muñeca y sensor colocado en el pecho ofrece la mejor tasa de precisión. Destaca también la sensibilidad calculada para los sensores colocados en la oreja y en el brazo.

- *Intensidad alta.* Los sensores colocados en brazo y rodilla representan bastante bien a este grupo de actividades pero es el sensor colocado en la oreja el que ofrece los mejores resultados.

Pese a que la cintura y el pecho es uno de los lugares más utilizados para la colocación del acelerómetros (Heil et al. 2012) (Murphy 2009), se puede comprobar la validez e interés que presenta la colocación del acelerómetro en la muñeca. Debe mencionarse el estudio realizado en (Atallah and Yang 2009) dónde se corrobora la independencia de la localización del acelerómetro para la medida de la actividad física.

En este trabajo, se propone el uso de un acelerómetro integrado en una prenda de vestir o complemento comúnmente vestido por el usuario (como relojes) para la monitorización de personas mayores. Según el estudio que se realizó en la sección 4.2 sobre las ventajas e inconvenientes de sensores vestibles y sensores ambientales, el uso de sensores vestibles integrados en prendas de vestir o complementos permite aprovecharse de las ventajas de su uso evitando sus inconvenientes (Steele et al. 2009). Así, en este trabajo se consigue un compromiso entre la exactitud en las medidas y la reluctancia de los usuarios a la vez que se propone un sistema multiusuario que podrá funcionar dentro y fuera del hogar.

4.4.3. MÉTODO DE PROCESADO (RF-2)

4.4.3.1. INTRODUCCIÓN

Las limitaciones anteriormente expuestas en el uso de acelerómetros para la medida de la actividad física (consumo eléctrico e inexistencia de ecuaciones estándar de transformación) son resueltas en este trabajo mediante el diseño de un nuevo método de procesado de la aceleración.

Este nuevo método de procesado permite la transformación de la aceleración mediante un nuevo estimador de actividad física que reduce el consumo eléctrico y la complejidad computacional frente a los estimadores existentes y cuya validez se ha demostrado experimentalmente. Este estimador funciona de forma automática adquiriendo medidas de forma pasiva (sin interacción del usuario) y realizando su transformación a actividad física.

El método de procesado propuesto para la monitorización, al no requerir interacción del usuario en su funcionamiento, supone una ventaja clara sobre los dos principales sistemas de monitorización existentes en la actualidad: “botón del pánico” y monitorización de variables médicas. Debido a las características y habilidades de las

personas mayores en el ámbito tecnológico, el manejo de dispositivos de monitorización es complejo. El sistema propuesto funciona de forma transparente sin que el usuario tenga que realizar ningún tipo de interacción con el mismo.

La aceleración medida por un acelerómetro vestido por un ser humano es debida principalmente a tres componentes principales (Bouten et al. 1997):

1. *Aceleración gravitacional.* Aceleración que posee todo cuerpo que se encuentra sobre la superficie terrestre. La aceleración gravitacional es la principal componente de la aceleración en en reposo.
2. *Aceleración cinemática.* Es la aceleración ocasionada por las actividades realizadas por el usuario. Esta componente describe el movimiento realizado.
3. *Vibraciones.* Esta componente engloba tanto las vibraciones ocasionadas por elementos externos como las relacionadas con el propio movimiento del sensor y el ruido que éste introduce al sistema.

De las componentes anteriores, solamente las dos primeras están relacionadas con el movimiento del cuerpo humano. La aceleración gravitacional variará entre los ejes de medida del acelerómetro y entre $+1\text{ g}$ y -1 g dependiendo de la orientación del dispositivo y por tanto, de la postura del usuario. Generalmente esta componente es conocida como la componente gravitacional de la aceleración.

La Figura 4.4 muestra la variación de la componente gravitacional en función de la posición del acelerómetro. Según sea su posición sobre la superficie, la componente gravitacional será medida en ejes diferentes del acelerómetro. La segunda de las componentes, conocida como la componente cinética de la aceleración, dependerá de la actividad realizada por el cuerpo en el que se encuentra el acelerómetro y está intrínsecamente relacionada con la actividad realizada por el usuario.

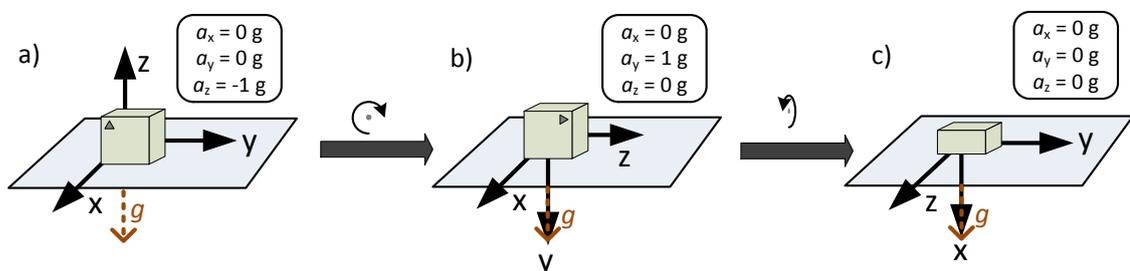


Figura 4.4. Aceleración gravitacional de un cuerpo en reposo

Existen varias características acelerométricas que pueden ser utilizadas para la monitorización de usuarios (Lai, Palaniswami, and Begg 2011) (Atallah et al. 2010) (Avci and Bosch 2010). Parte de este tipo de características pueden verse resumidas en la Tabla 4.15. En esta tabla se hace una división en características temporales, características frecuenciales y características de tipo heurístico.

Las características mostradas en la Tabla 4.15 son utilizadas principalmente para el reconocimiento de actividad. A partir de la actividad identificada, se realiza la transformación a actividad física utilizando las tablas presentadas en (Ainsworth et al. 2011). La mayoría de estos trabajos se basan en desarrollos experimentales y no poseen una validez científica que las relacione con la actividad física. Además, el error cometido en el reconocimiento de actividad se acarrea para la estimación de la actividad física.

De entre todas las medidas acelerométricas existentes para la medida de la actividad física, destaca el área de la variable de la aceleración (*signal magnitude area, SMA*) ampliamente utilizada en trabajos científicos como (Curone et al. 2010) o (Karantonis et al. 2006). Su relación con las calorías consumidas y por tanto su validez para la medida de la actividad física fue por primera vez demostrada en (Bouten et al. 1997) y desde entonces utilizada en numerosos trabajos.

La derivada de la aceleración o *jerk*, es una medida interesante para la actividad física ya que representa la tasa de cambio de la aceleración. Esta medida permite caracterizar la forma en la que varía la componente cinemática de la aceleración de un cuerpo identificando sus variaciones.

En este trabajo se propone un nuevo estimador de actividad física basado en la derivada de la aceleración y que es validado directamente estudiando su relación con el estimador SMA. Al realizar la medida del comportamiento en términos de actividad física sin realizar un reconocimiento previo de actividad se consiguen mejores y más eficientes resultados.

Tabla 4.15. Características acelerométricas para la medida de la actividad física.

Tipo	Característica	Descripción
Temporal	Variable en bruto	Uso de la aceleración tal y como es medida por el acelerómetro. No se tiene en cuenta la descomposición en componentes
	Morfología de la variable	A través de la forma de la variable se realiza un reconocimiento de actividad en el que se basa la medida de la actividad física
	Derivadas de la variable	Cálculo de las derivadas de la aceleración: <i>jerk</i> (primera derivada) y <i>snap</i> (segunda derivada). Representan la sobre-aceleración y la tasa de variación de la sobre-aceleración respectivamente
	Características estadísticas de la variable	A través de ventanas deslizantes se realiza el cálculo de características estadísticas de la aceleración (media, varianza, desviación, curtosis...) que posteriormente son utilizadas para la medida de actividad
Frecuencial	Componentes de la variable	Mediante técnicas de filtrado se descompone la aceleración en sus componentes frecuenciales diferenciando la componente gravitacional de la componente cinemática
	Estudio espectral de la variable	Descomposición en las principales componentes frecuenciales de la variable y estudio de sus características (frecuencia, energía...)
	Transformada Wavelet	Transformada wavelet de la aceleración que permite su representación mediante wavelets
Heurísticas	Área de la variable	Área encerrada bajo la curva de la aceleración. Se calcula mediante la integral de la componente cinemática de la aceleración (SMA)
	Coeficientes de correlación	Estudio de la relación entre componentes de la aceleración (diferentes ejes o diferentes dispositivos colocados sobre el cuerpo) mediante el análisis de la correlación entre variables

4.4.3.2. SIGNAL MAGNITUDE AREA (SMA)

La Figura 4.5 muestra cómo la SMA es calculada a partir de la variable de la aceleración. La SMA representa para cada muestra el área en valor absoluto envuelto por la variable de la aceleración.

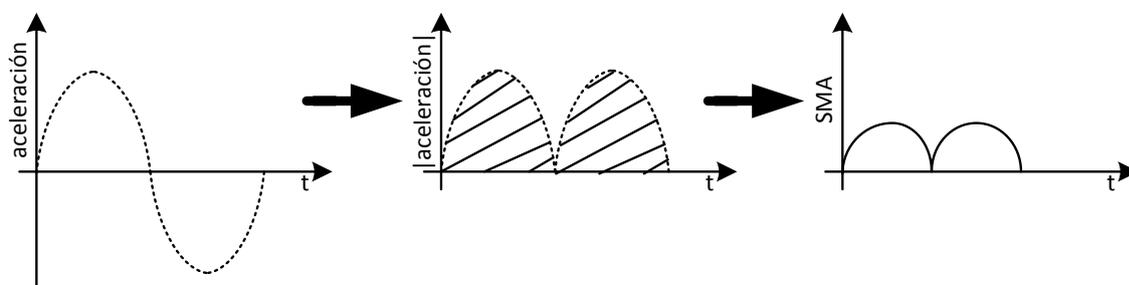


Figura 4.5. Representación gráfica del cálculo de la SMA.

Como ya se mencionó en la sección 4.4.2, para realizar una correcta caracterización de la actividad física es necesario que el acelerómetro sea muestreado al menos a 40 Hz, ya que la mayoría de las componentes espectrales ocasionadas por el movimiento del cuerpo humano (componente cinemática de la aceleración) se encuentran en el rango de 0,3 Hz a 20 Hz. Tal y como se representa en la Figura 4.6, para el cálculo de la SMA el acelerómetro generalmente se muestrea a 50 Hz. Siguiendo el diagrama de bloques de la Figura 4.6, el cálculo de la SMA puede dividirse en las siguientes dos etapas:

- *Filtrado de mediana.* Su objetivo es realizar una eliminación del posible ruido que haya contaminado la variable. Con este tipo de filtrado se eliminan principalmente las componentes de la aceleración debidas a vibraciones externas y vibraciones propias del sensor. En función de la aplicación (y su susceptibilidad al ruido) se seleccionará el tamaño del filtro para conseguir eliminar toda la posible información no deseada.
- *Filtrado paso-bajo.* Mediante este tipo de filtro se hace una separación espectral de las frecuencias que componen la aceleración. La idea es separar la componente gravitacional (componente continua) de la componente cinemática (componente de alta frecuencia) de la aceleración. La frecuencia de corte de este filtro debe encontrarse en torno a 0,3 Hz ya que las componentes superiores a esa frecuencia son las que originan principalmente la actividad física (Plasqui et al. 2013). En (Curone et al. 2010) se propone el uso de un filtro paso bajo elíptico de tipo IIR y de tercer orden (frecuencia de corte de 0,3 Hz, 0,1 dB de rizado pasabanda y una atenuación de -100 dB). La salida de este filtrado paso-bajo será la componente gravitacional.

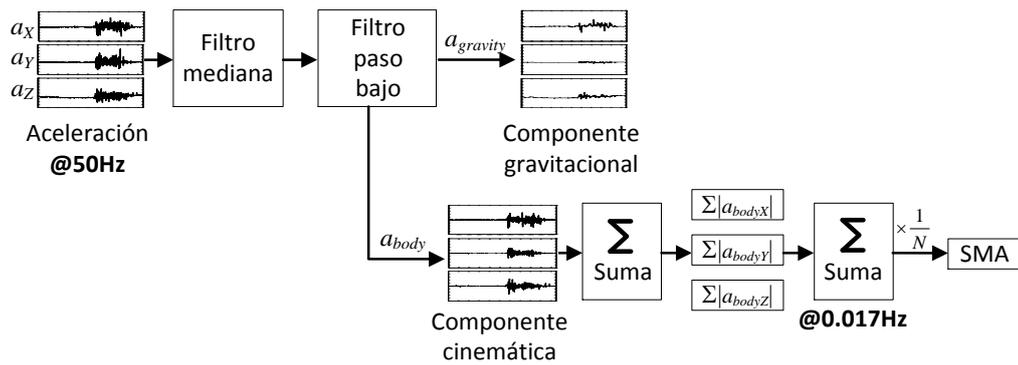


Figura 4.6. Diagrama de bloques para el cálculo de la SMA

La componente cinemática, ocasionada por la actividad física y obtenida tras el filtrado paso-bajo, es la utilizada para el cálculo de la SMA siguiendo la Ecuación 4.1. Generalmente la SMA es calculada cada minuto (0,017 Hz; $T = 1$ minuto; $N = 3000$ muestras) caracterizando así la actividad física realizada durante ese minuto en el que se realiza el cálculo. Teniendo en cuenta la expresión de la Ecuación 4.1, la SMA puede ser también interpretada directamente como la suma de las aceleraciones observadas en un determinado intervalo de tiempo (denominado *epoch* y con valor de 1 minuto en la Figura 4.6) en la componente cinemática de la aceleración (Heil et al. 2012). Así, en algunas referencias se habla de sumatorio de la aceleración en lugar de SMA. Generalmente, la unidad de medida utilizada es *count* que hace referencia a los valores obtenidos directamente de la cuantificación analógica-digital que tiene lugar en el acelerómetro (Taraldsen et al. 2012) (Chen and Bassett 2005). El uso de las *cuentas* o *counts* simplifica los cálculos sin necesidad de realizar transformaciones entre los valores leídos directamente del sensor y los valores reales de aceleración.

$$SMA = \frac{1}{N} \left(\sum_0^N |a_{bodyx}| + \sum_0^N |a_{bodyy}| + \sum_0^N |a_{bodyz}| \right)$$

Ecuación 4.1

La utilización de la SMA para la medida de la actividad física presenta dos inconvenientes principales que han de ser tenidos en cuenta en su utilización y que han sido ampliamente identificados en varios trabajos como (Murphy 2009) (Heil et al. 2012) (Ward et al. 2005) (Chen and Bassett 2005):

- *Coste computacional.* El algoritmo implicado en el cálculo de la SMA (ver Figura 4.6) presenta una alta complejidad al ser necesario realizar varios tipos de filtrado (filtrado de mediana y filtrado paso-bajo). La implementación computacional del cálculo no es trivial e implica que el dispositivo en el que el

acelerómetro se encuentra integrado tenga capacidad computacional suficiente para su implementación.

- *Tiempo de vida.* Al ser necesario trabajar a una tasa de muestreo tan alta (> 40 Hz) el dispositivo ha de estar continuamente muestreando al acelerómetro y en general, el tiempo de vida de los dispositivos de medida de actividad es bajo. Esta limitación ya había sido identificada en la sección 4.4.2 como uno de los principales inconvenientes en el uso de acelerómetro como sensores de medida. Un aumento de tamaño del sistema de alimentación podría permitir aumentar el tiempo de vida pero se reduciría la portabilidad y se incrementaría la invasividad en la monitorización.

En este trabajo se propone el diseño de un nuevo estimador de actividad física basado en la aceleración que permita la medida de la actividad física y que resuelva los principales inconvenientes del uso de acelerómetros y de la SMA: disminución de la complejidad computacional del algoritmo, aumento del tiempo de vida del dispositivo y uso eficiente de la memoria del dispositivo.

El estimador de actividad física propuesto en este trabajo (descrito en la sección 4.4.3.3) disminuye la complejidad computacional de la medida respecto al uso de la SMA, aumenta consecuentemente el tiempo de vida del dispositivo y consigue utilizar de forma más eficiente la capacidad de memoria.

4.4.3.3. JERK-BASED INFORMATION MAGNITUDE (JIM)

Dentro de las características acelerométricas descritas en la Tabla 4.15 se encuentra la derivada de la aceleración o *jerk* que ha sido utilizada en algunos trabajos como (Gyórbíró, Fábán, and Hományi 2008). El *jerk* se define como la tasa temporal de variación de la aceleración y representa la sobre-aceleración que experimenta un cuerpo cuando se encuentra en movimiento (Zatsiorsky 2002). Tal y como se muestra en la Ecuación 4.2, el *jerk* puede ser definido como la primera derivada de la aceleración respecto al tiempo y por tanto será equivalente a la segunda derivada de la velocidad o a la tercera derivada temporal del desplazamiento. Suele nombrarse mediante la letra \vec{j} .

$$\vec{j} = \frac{\partial \vec{a}}{\partial t} = \frac{\partial^2 \vec{v}}{\partial t^2} = \frac{\partial^3 \vec{r}}{\partial t^3}$$

Ecuación 4.2

Partiendo de los principales problemas en el uso de acelerómetros y la SMA para la medida de la actividad física (coste computacional y tiempo de vida), en este trabajo se propone un nuevo estimador de actividad física basado en el cálculo de la derivada de la aceleración y que fue descrito por primera vez en (Carús et al. 2012). A esta nueva característica acelerométrica se la denomina JIM (*jerk-based information magnitude*) y se basa en el diseño de un algoritmo de poca complejidad, eficiente en energía y que requiera poco uso de memoria para su almacenamiento.

- *Coste computacional.* Con el objetivo de reducir el coste computacional, el algoritmo utilizado para el cálculo de JIM reduce las etapas de filtrado necesarias para el cálculo de la SMA.
- *Tiempo de vida.* El estimador utilizado para la medida influye principalmente en el tiempo de vida a través de la tasa de muestreo del acelerómetro. Si se reduce la frecuencia a la que se adquieren los datos, se conseguirá reducir el gasto energético que implica el uso de este elemento. La Tabla 4.16 muestra la variación en el consumo eléctrico que se consigue cuando se varía la tasa de muestreo del acelerómetro. Los resultados se basan en los datos obtenidos en las hojas de características de tres acelerómetros comerciales ampliamente extendidos e integrados en numerosos dispositivos en la actualidad (CMA3000-D01, ADXL345 y MMA8450Q).

Tabla 4.16. Consumo eléctrico de varios acelerómetros comerciales

Acelerómetro	Consumo a diferentes frecuencias		Ratio
	50 Hz	1Hz	
CMA3000-D01 (VTI Technologies 2009)	50 μ A	9 μ A	5,55
ADXL345 (Analog Devices 2013)	100 μ A	40 μ A	2,50
MMA8450Q (Freescale 2012)	42 μ A	27 μ A	1,55

Siguiendo los dos puntos anteriores (coste computacional y tiempo de vida), el estimador de actividad propuesto en este trabajo (JIM) elimina las etapas de filtrado necesarias para el cálculo de la SMA y reduce la frecuencia de adquisición del acelerómetro de 50 Hz ($T = 20$ milisegundos) a 1 Hz ($T = 1$ segundo). Por tanto, el estimador propuesto reduce la complejidad computacional y el tiempo de vida de la batería en el cálculo de la actividad física.

La Ecuación 4.3 muestra la expresión matemática a través de la cual se define el JIM para un conjunto de N muestras. Para cada uno de los ejes en los que realiza la medida de la aceleración (x, y, z), se calcula su derivada. Esta derivada, definida como la tasa temporal de variación de la aceleración, se calcula mediante las diferencias temporales que indican como la aceleración varía con el tiempo. Para cada uno de los ejes, el *jerk* calculado para las N muestras es sumado y promediado para la obtención del valor de JIM. Este valor define la actividad física realizada durante el intervalo temporal formado por las N muestras.

$$JIM = \frac{1}{N} \sum_{i=1}^N [|Jerk(i)|_x + |Jerk(i)|_y + |Jerk(i)|_z]$$

Ecuación 4.3

La Figura 4.7 muestra el diagrama de bloques que definen el cálculo del JIM. La tasa de adquisición del acelerómetro decrece de 50 Hz en el cálculo de la SMA a 1 Hz en el cálculo del JIM. Para cada una de las muestras de aceleración adquiridas se calcula la primera derivada de la aceleración obteniendo para cada eje el valor absoluto de la tasa de variación temporal. Este valor será calculado cada segundo (1 Hz). Cada minuto (0,017 Hz), los 60 valores ($N = 60$) obtenidos en el bloque del cálculo de *jerk* serán sumados y promediados para la obtención del JIM.

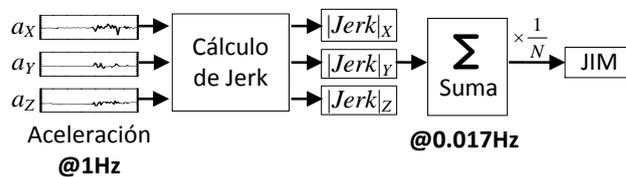


Figura 4.7. Diagrama de bloques para el cálculo del JIM

El pseudocódigo cuya implementación realiza el cálculo del JIM se muestra en la Figura 4.8. Este código deberá ser implementado en el dispositivo con acelerómetro integrado que desee ser utilizado para la medida de la actividad física.

Cada segundo, se muestrea cada uno de los ejes de medida del acelerómetro (método *leerAcelerómetroTriaxial*) obteniendo la aceleración en cada eje del movimiento (a_x, a_y, a_z). Estas medidas son utilizadas para calcular la derivada (*calcularJerk*) respecto al valor anterior medido en cada eje ($Jerk_x, Jerk_y, Jerk_z$). Cada minuto, las derivadas calculadas son sumadas para siguiendo la expresión de la Ecuación 4.3 calcular el JIM.

```

Integer i = 0;
Integer j = 0;
Integer[60] Jerkx, Jerky, Jerkz;
Integer[] JIM;
Every 1 second do:
    Double ax, ay, az = leerAcelerómetroTriaxial();
    Jerkx[i], Jerky[i], Jerkz[i] = calcularJerk(ax, ay, az);
    i = i + 1;
    if (i == 60) do: // cada 60 muestras = cada 60 segundos
        JIM[j] = JIM(Jerkx, Jerky, Jerkz); // ver Ecuación 4.3
        j = j + 1;
        i = 0;
    End do;
End do;

```

Figura 4.8. Pseudocódigo para el cálculo del JIM

En comparación con el cálculo de la SMA, el cálculo de JIM implica un muestreo del acelerómetro de solo 1Hz. En ambos casos la actividad física se calcula cada minuto pero mientras que en la SMA se utilizan 3000 muestras, en el cálculo del JIM solamente intervienen 60 muestras. La complejidad del algoritmo baja al no ser necesario ningún tipo de filtrado y trabajar con una menor cantidad de datos.

El nuevo estimador de actividad (JIM) propuesto en este trabajo pretende mejorar la eficiencia del estimador SMA, ampliamente utilizado en la medida de la actividad física, tanto en términos de complejidad como en tiempo de vida de la batería. Para su uso como estimador de actividad es necesario llevar a cabo una etapa experimental de validación.

4.4.4. VALIDACIÓN DEL MÉTODO PROPUESTO

4.4.4.4. INTRODUCCIÓN

La validación del método de medida propuesto se basa en el análisis de la validez de JIM como estimador de la actividad física. Esta validación parte de la relación lineal existente entre la SMA y la actividad física. Esta relación fue probada en (Bouten et al. 1997) y está ampliamente aceptada por la comunidad científica. En este trabajo se validará al JIM como estimador de la actividad física estudiando su relación con el estimador SMA.

La relación entre el JIM y la SMA se ha estudiado experimentalmente utilizando como sensor un dispositivo vestibular con apariencia de reloj y el procedimiento de experimentación descrito en el Apéndice G. Este dispositivo vestibular ha sido desarrollado por Texas Instruments y es conocido como TI eZ430-Chronos (Texas

Instruments 2014). El algoritmo mostrado en la Figura 4.8 para el cálculo del JIM fue implementado junto con el algoritmo necesario para el cálculo de la SMA siguiendo el diagrama de bloques de la Figura 4.6 y la expresión matemática de la Ecuación 4.1.

Para la validación del método de medida propuesto, 39 usuarios (21 hombres y 18 mujeres) pertenecientes al segmento de edad adulto (de 20 años a 60 años) vistieron el dispositivo con apariencia de reloj durante las actividades de la vida diaria y durante al menos 24 horas (media de 30,17 horas y $\sigma = 2,24$). La media de edad de los usuarios fue de 37,44 años ($\sigma = 11,19$) con un índice medio de masa corporal (BMI) de 24,56 ($\sigma = 3,97$). La Tabla 4.17 muestra el resumen de las características que definen a los usuarios involucrados en el sistema y que son descritas en detalle en el Apéndice D. En esa sección del apéndice se identifica para cada usuario su perfil de actividad con el objetivo de mostrar la variabilidad de perfiles utilizados en la validación.

Tabla 4.17. Usuarios involucrados en la validación del método de medida

	Media	σ
Edad	37,44 años	11,19 años
Peso	69,71 Kg	12,81 Kg
Estatura	1,68 m	0,08 m
Índice de masa corporal	24,56	3,97
Duración del experimento	30,17 horas	2,24 horas

Pese a que el sistema de monitorización final está pensado para su uso por personas mayores, la validación del JIM como estimador de la actividad física tiene lugar con usuarios del segmento adulto (de entre 20 y 60 años). Este tipo de usuarios realizan actividades con una mayor variación de intensidad que las personas mayores tal y como se indica en varias referencias como (Murphy 2009). Una validación solo con personas mayores podría sesgar los resultados al validar solamente el JIM como estimador de actividades de baja intensidad. No obstante, el diseño de todo el método de medida se ha realizado teniendo en cuenta los requisitos necesarios para su uso por personas mayores.

La relación entre la SMA y el JIM fue estudiada siguiendo la directrices dadas en (Welk, McClain, and Ainsworth 2012) para la evaluación de estimadores de actividad. Partiendo de la validez de SMA como estimador de la actividad física, se valida el estimador JIM mediante el análisis de la relación entre ambos. El análisis entre ambos estimadores de actividad física se hace en base a tres indicadores diferentes: asociación, sesgo y diferencias entre individuos. A su vez, se realiza una comparación de sus características computacionales.

4.4.4.1. ASOCIACIÓN ENTRE VARIABLES

Para la validación del JIM como estimador de la actividad física, se estudia en primer lugar la asociación existente entre este nuevo estimador propuesto y el estimador SMA a través del coeficiente de correlación de Spearman. Este coeficiente permite analizar la relación lineal existente entre dos variables aleatorias cuantitativas.

El estudio de la correlación entre dos variables indica la fuerza de la relación lineal y proporcional entre ellas. Se considera que dos variables están correlacionadas cuando los valores de una de ellas varían y se registran variaciones homónimas en la otra. Es decir, incrementos en una variable conllevan incrementos en la otra y viceversa.

La Figura 4.9 muestra el histograma del coeficiente de correlación de Spearman entre el estimador de actividad JIM y el estimador SMA para el conjunto de usuarios participantes en el experimento. Los resultados mostrados prueban la fuerte asociación existente entre ambos estimadores con una media en el coeficiente de correlación obtenido de 0,95 ($\sigma = 0,03$; p-value < 0,05).

Para cada uno de los usuarios del experimento se ha calculado la correlación entre la actividad medida con el estimador JIM y con el estimador SMA durante el periodo de tiempo que transcurrió durante el experimento (duración mínima de 24 horas con 30,17 horas de media). Todos estos índices de correlación pueden ser consultados para cada usuario en el Apéndice D y son los valores a partir de los cuales se construyó el histograma de la Figura 4.9.

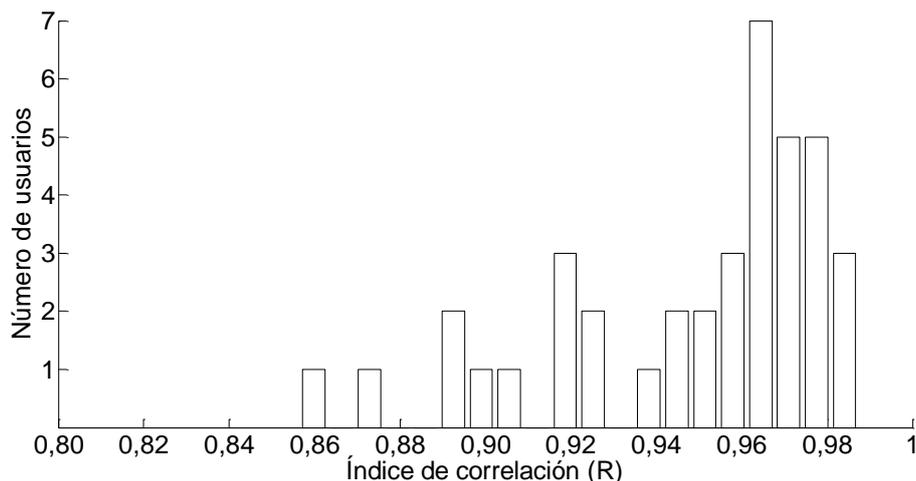


Figura 4.9. Histograma de la correlación entre el JIM y la SMA

4.4.4.2. ANÁLISIS DEL SESGO

Este análisis permite determinar si la estimación es consistente para todos los rangos de actividad medidos. El análisis del sesgo permite la detección de diferencias en la estimación de la actividad física en función del nivel de actividad realizado. El gráfico Bland-Altman es una de las herramientas más extendidas y aceptadas para el análisis del sesgo entre dos medidas (Dewitte and Fierens 2002).

La Figura 4.10 muestra el gráfico de Bland-Altman para las dos estimaciones de actividad física. Este gráfico se construye representando en el eje x la diferencia obtenida en la estimación media de actividad para cada usuario durante el experimento ($JIM_{media} - SMA_{media}$) y en el eje y la media de las dos estimaciones ($(SMA_{media} + JIM_{media})/2$). En el Apéndice D se muestra para cada usuario el cálculo total para los dos estimadores (JIM_{total} y SMA_{total}) y la duración del experimento. El resultado del cociente será respectivamente JIM_{media} y SMA_{media} .

La línea horizontal negra representada en la Figura 4.10 muestra la media de la diferencia (2,83 counts) y las líneas grises muestran el valor correspondiente al doble de la desviación estándar de la misma ($\pm 1,24$ counts). Al encontrarse todos los valores representados dentro del intervalo fijado por las líneas, se prueba la consistencia en las medidas y cómo la diferencia en la estimación no se ve influenciada por el valor de medida (Welk et al. 2012).

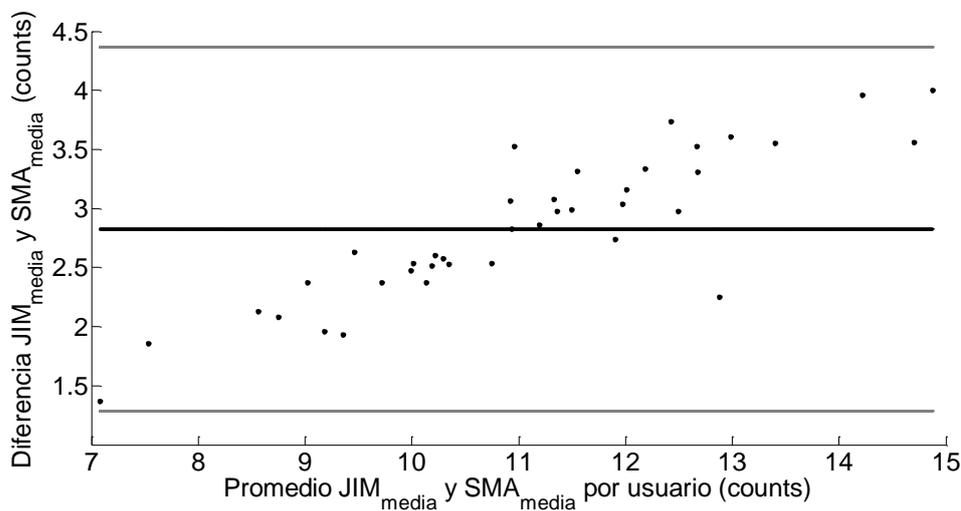


Figura 4.10. Gráfico de Bland-Altman entre JIM y SMA media por usuario

4.4.4.3. DETECCIÓN DE DIFERENCIAS ENTRE INDIVIDUOS

Es necesario analizar posibles diferencias entre grupos de usuarios. Un análisis estadístico de las medidas tomadas por grupo (edad, peso, estatura) permitiría detectar características especiales de la medida que no fueran tenidas en cuenta en un primer momento.

Para estudiar el comportamiento individual del JIM como estimador de actividad, se calcula para cada usuario la diferencia minuto a minuto entre el estimador SMA y el estimador JIM (los valores medios por usuario se muestran en el Apéndice D). La Ecuación 4.4 muestra el cálculo realizado para cada una de las muestras t de medida. Este cálculo se realizará sobre las medidas tomadas en los 39 experimentos.

$$\text{Diferencia de estimación}(t) = JIM(t) - SMA(t)$$

Ecuación 4.4

Estas diferencias son representadas en la Figura 4.11. Mediante un diagrama de cajas se muestran las diferencias calculadas para cada usuario. El diagrama se ha construido indicando la mediana con una línea recta y estableciendo los percentiles 25 (Q1) y 75 (Q3) de la muestra como límites. Para todos los usuarios se aprecia una diferencia similar en la estimación de media 2,83 counts y desviación estándar 0,62.

A la vista de los resultados, se puede determinar que la diferencia entre la SMA y el JIM sigue un patrón similar para todos los usuarios. Así, no pueden establecerse ningún tipo de diferencia en la estimación dependiendo de las características de los usuarios (edad, sexo, índice de masa corporal...) ni de su perfil de actividad.

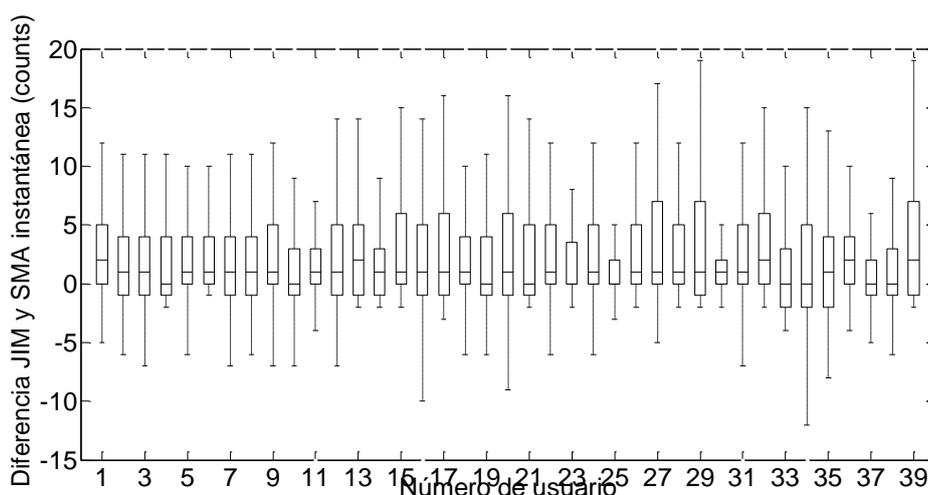


Figura 4.11. Diferencia estimada entre el valor de JIM y SMA para cada usuario

4.4.4.4. CARACTERÍSTICAS COMPUTACIONALES

Con el estudio realizado de validación, se ha probado que el estimador de actividad JIM propuesto para la medida de la actividad física es un estimador tan válido como la SMA. El estudio realizado ha probado la asociación existente entre ambas medidas en términos de correlación, diferencia de grupo, diferencia individual y sesgo. Así, el JIM es un estimador de actividad física tan válido como el estimador SMA y que tal y como se analizará a continuación, presenta mejores características computacionales.

La Tabla 4.18 muestra las principales características computacionales para el cálculo de la SMA y para el cálculo del JIM utilizando el dispositivo vestible con apariencia de reloj (TI eZ430-Chronos) utilizado en la experimentación. La principal característica del uso de JIM reside en el decremento de la tasa de muestreo del acelerómetro que hace que se incremente alrededor de 5 veces el tiempo de vida del dispositivo.

Tabla 4.18. Comparación el cálculo de la SMA y del JIM

Parámetro	SMA	JIM	Relación
Frecuencia de muestreo	50 Hz	1 Hz	50:1
Frecuencia de almacenamiento	0,017 Hz	0,017 Hz	1:1
Tamaño de muestra	2 bytes	2 bytes	115:110
Tamaño de código	1843 bytes	1589 bytes	50:9
Consumo del acelerómetro	50 μ A	9 μ A	50:9

El uso de JIM permite aprovechar las ventajas del uso de acelerómetros vestibles para la medida de la actividad física resolviendo simultáneamente los inconvenientes relacionados con el uso de SMA y en general de acelerómetros: coste computacional, tiempo de vida y capacidad de memoria. La Figura 4.12 muestra una representación gráfica de los resultados obtenidos en la comparación entre la SMA y el JIM con el dispositivo TI eZ430-Chronos. Al trabajar a una frecuencia de muestreo menor, con una menor necesidad de almacenamiento y un menor consumo eléctrico se consigue un decremento del coste computacional del método de medida, un aumento del tiempo de vida y un decremento de la capacidad de memoria requerida.

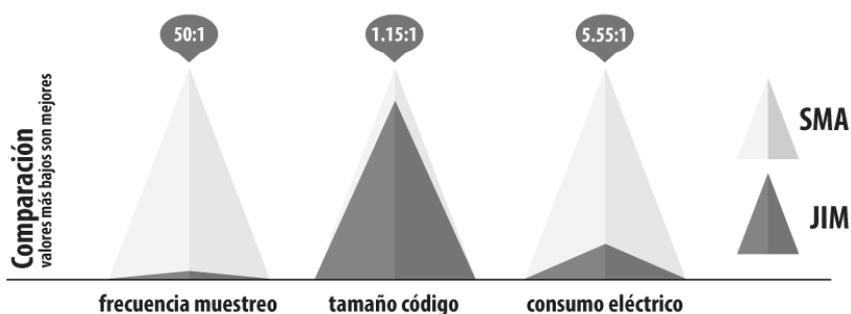


Figura 4.12. Comparación entre la implementación de SMA y JIM

4.4.5. RESUMEN

La Figura 4.13 muestra el diagrama de bloques del método de medida propuesto de acuerdo a los bloques expuestos en la Figura 2.4.

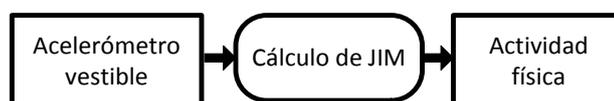


Figura 4.13. Diagrama de bloques del método de medida propuesto

A través de un acelerómetro vestible se realiza una adquisición pasiva (sin interacción del usuario) de la aceleración experimentada. Esta aceleración es transformada a actividad física (variable de salida del método de medida) mediante el estimador JIM propuesto y enunciado en la Ecuación 4.3. Este estimador requiere el muestreo del acelerómetro a una tasa de 1 Hz realizando el cálculo de la actividad física cada minuto con $N = 60$ muestras de aceleración. La Tabla 4.19 muestra el resumen de los parámetros implicados en el método propuesto.

Tabla 4.19. Resumen de los parámetros implicados en el método de medida

Parámetro	Valor propuesto	Descripción
T	1 segundo	Tiempo de muestreo de la aceleración
N	60	Número de muestras implicadas en el cálculo de JIM

La Tabla 4.20 parte de los requisitos funcionales que debe cumplir un método de medida del comportamiento y expone las decisiones tomadas en este capítulo para su diseño. Se puede comprobar cómo se cumplen todos los requisitos enunciados en el Capítulo 2.

Tabla 4.20. Requisitos funcionales del método de medida

Referencia	Requisito	Descripción
☑ RF-1	Sensor de medida	<i>Se utilizará un acelerómetro vestible integrado en una prenda vestida comúnmente por el usuario</i>
☑ RF-2	Método de procesado	<i>Partiendo de la aceleración medida por el acelerómetro se calculará el JIM mediante la Ecuación 4.3</i>
☑ RF-3	Magnitud de medida	<i>La magnitud seleccionada es la actividad física</i>

4.5. CONCLUSIONES

En vista de los sensores y magnitudes utilizados en el estado del arte para la monitorización del comportamiento, en este trabajo se ha optado por el uso de la actividad física por tratarse de una magnitud estándar. Esta magnitud posee una relación bidireccional con el estado de salud del usuario y ha sido utilizada principalmente con fines deportivos para la monitorización de personas.

De todos los sensores existentes para la medida de la actividad física destacan los acelerómetros por ser elementos sencillos, ligeros y ampliamente integrados en numerosos dispositivos. El área de la señal de la aceleración (*signal magnitude area*, SMA) es uno de los estimadores de actividad física más utilizados pero que no resuelve los principales inconvenientes en el uso de acelerómetros al presentar un coste computacional alto.

En este trabajo se propone el uso de la actividad física para la medida del comportamiento en personas mayores y se ha diseñado un nuevo método de medida que permite la adquisición de datos del usuario de forma transparente y maximizando la autonomía. Para ello, se ha propuesto un nuevo estimador de actividad que presenta mejores características computacionales que los existentes hasta el momento. Este nuevo estimador al que se ha denominado JIM (*jerk-based information magnitude*) se basa en la derivada de la aceleración para obtener una medida de la actividad física realizada por el usuario.

El uso del estimador JIM permite reducir el coste computacional del método de medida, aumentar el tiempo de vida del dispositivo y realizar un uso eficiente de su capacidad de almacenamiento. El uso del JIM como estimador de la actividad física ha sido validado experimentalmente en condiciones reales mediante su relación con la

SMA, estimador de actividad ampliamente aceptado. En la etapa de experimentación han participado 39 usuarios de diferentes perfiles de actividad y que fueron monitorizados durante más de 24 horas. La asociación entre JIM y SMA ha sido analizada a nivel de correlación (índice de correlación de Spearman de 0,95; p-value < 0,05), sesgo y diferencias entre individuos en la estimación. El análisis realizado ha demostrado la validez de JIM como estimador de la actividad física.

CAPÍTULO 5.

MÉTODO DE MODELADO:

IDENTIFICACIÓN DE RUTINA

5.1. INTRODUCCIÓN

Tras realizar la medida de la actividad física, es necesaria la existencia de un método de procesamiento que dote de inteligencia a la monitorización. Cuando el objetivo de la monitorización es la detección de anomalías, el modelado de las medidas se convierte en el principal objetivo de la etapa de procesamiento (Esling and Agon 2012). Al estar hablando del comportamiento del usuario a través de la medida de la actividad física, la extracción de un modelo implica la identificación de la rutina del usuario. Se entiende por rutina al conjunto de costumbres o hábitos que definen la actividad realizada por una persona en su día a día. Una rutina implica la realización de actividades de forma casi automática, sin razonamiento.

El objetivo de este capítulo es el “Diseño y validación de un método de modelado” que permita identificar la rutina del usuario a partir de la medida de su comportamiento utilizando la actividad física. La Figura 5.1 describe gráficamente este objetivo. Siguiendo las etapas descritas en la Figura 2.5 (acondicionamiento y procesado), en este capítulo se parte de la magnitud seleccionada en el Capítulo 4 (actividad física) y se diseña un método de acondicionamiento (RF-4) y un método de procesado (RF-5) que permitan la identificación de la rutina del usuario (modelo esperado, RF-6).

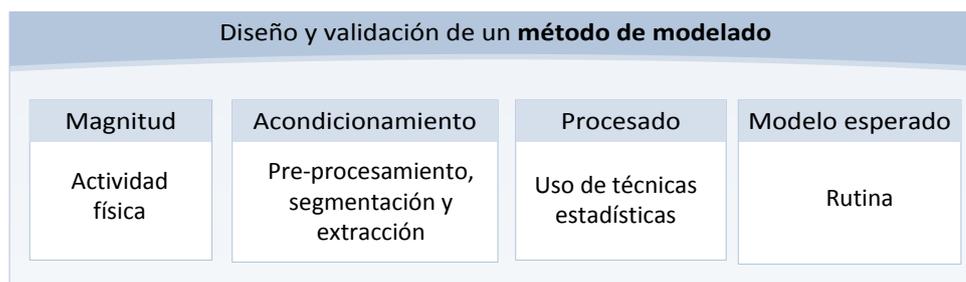


Figura 5.1. Objetivo principal del Capítulo 5

En la sección 5.2 se presentan las principales aproximaciones existentes en el estado del arte para cada una de las etapas que conforman el método de modelado. Se describen las principales técnicas existentes junto con sus ventajas e inconvenientes tanto para el acondicionamiento como para el procesado de las medidas. En la etapa de acondicionamiento (sección 5.2.2) destacan las técnicas de pre-procesamiento, segmentación y extracción de características. El objetivo principal de esta etapa es garantizar la calidad en los datos que llegan al método de procesado. Las ventajas e inconvenientes de cada técnica y su aplicabilidad para el procesado del comportamiento humano son estudiados y analizados en este capítulo.

Dentro de los métodos para el procesado del comportamiento (sección 5.2.3) destacan los métodos basados en modelado teórico, el uso de técnicas de reconocimiento de actividad y el uso de métodos de aprendizaje automático. Este último tipo de métodos destacan por ser capaces de identificar de forma automática y adaptativa la rutina del usuario (requisitos de este trabajo) a partir de las medidas de actividad tomadas. El resto de alternativas requieren etapas previas de entrenamiento y la adición de conocimiento experto en el procesado.

En la última parte de este capítulo (sección 5.3), se presenta el método propuesto y las técnicas utilizadas en cada etapa. La salida del método es una variable que modela la rutina del usuario y que es construida a partir del modelado de las medidas de actividad física tomadas. El método de modelado propuesto permite a través de técnicas estadísticas la identificación de la rutina del usuario sin realizar ninguna asunción teórica y basándose únicamente en sus medidas históricas. Esto permite que el método propuesto se adapte de forma automática a los posibles cambios que pueda realizar el usuario en su rutina de forma natural y sin necesidad de reconfiguración o redefinición del método. Esta característica es señalada en algunos trabajos como (Estevez-Tapiador et al. 2004) o (Esling and Agon 2012) como uno de los principales retos de un método de procesado del comportamiento.

5.2. ESTADO DEL ARTE:

MODELADO DEL COMPORTAMIENTO

5.2.1. INTRODUCCIÓN

Un método para el modelado del comportamiento humano se divide en varios trabajos como (Avci and Bosch 2010) y (Biswas et al. 2010) en las 4 fases que se muestran en la Figura 5.2: pre-procesamiento de las medidas, segmentación, extracción de características y modelado. Estas fases se alinean con los bloques mostrados en la Figura 5.1 interpretando el método de acondicionamiento como la unión de las fases de pre-procesamiento, segmentación y extracción de características.

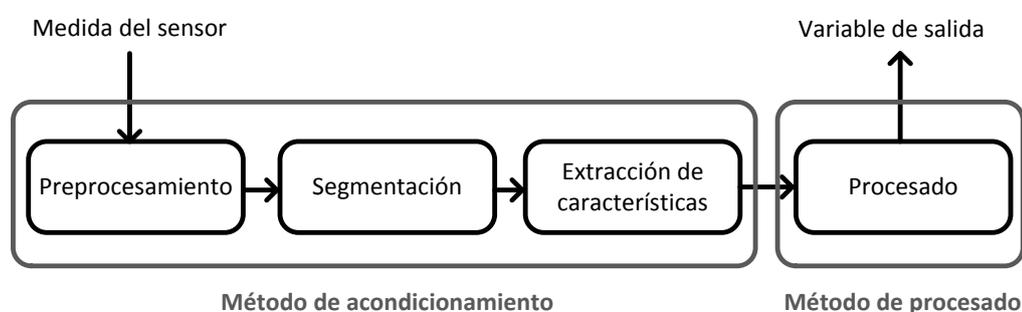


Figura 5.2. Etapas de un método de modelado del comportamiento

- *Pre-procesamiento*. La finalidad de esta etapa es garantizar la calidad de los datos con el objetivo de maximizar la fiabilidad del método de modelado.
- *Segmentación*. El principal objetivo de esta etapa es la obtención de segmentos de datos de menor tamaño. Esta etapa permite simplificar y mejorar los resultados de los algoritmos a aplicar en el resto de etapas.
- *Extracción de características*. Las técnicas de extracción de características permiten caracterizar la actividad física de forma efectiva. Destacan las características temporales (media, varianza), espectrales (energía espectral, coeficientes wavelets) y características propias del dominio (número de pasos, pulsación sanguínea).
- *Procesado*. A partir de las características extraídas se realizará un procesado del comportamiento con el objetivo de obtener un modelo que identifique la rutina del usuario.

Para cada una de las fases descritas, se muestran a continuación las principales tendencias encontradas en el estado del arte.

5.2.2. MÉTODO DE ACONDICIONAMIENTO

5.2.2.1. PRE-PROCESAMIENTO

El objetivo de la fase de pre-procesamiento es la preparación y manipulación de los datos para que puedan ser manejados o interpretados de forma más sencilla (Pyle 1999). En varios trabajos como (Davis and Clark 2011), (Kasteren et al. 2010) y (Aztiria et al. 2010) se señala la etapa de pre-procesamiento como una de las etapas más importantes en un método de modelado y detección de anomalías ya que permite la eliminación de características no deseadas y no esperadas en los datos y que podrían interferir en la correcta detección y funcionamiento del método.

La Figura 5.3, adaptada de (Radhakrishna et al. 2012), describe los siete objetivos principales de la fase de pre-procesamiento. Los datos a la salida de esta fase han de ser válidos, consistentes, objetivos, generales, útiles, completos, relevantes e íntegros.

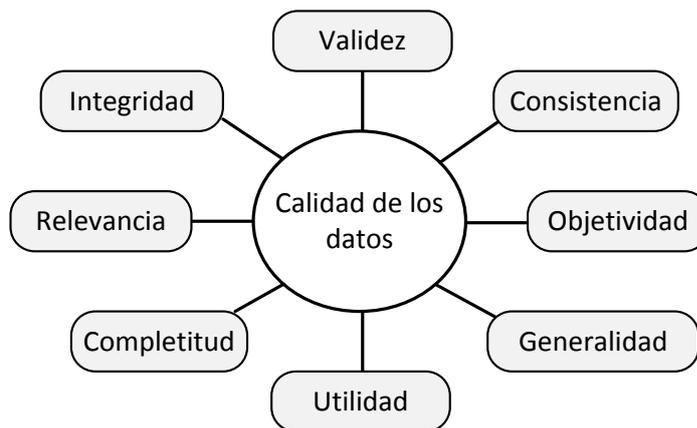


Figura 5.3. Características que garantizan la calidad de los datos

El pre-procesamiento de datos engloba todas aquellas técnicas de análisis que permiten mejorar la calidad de un conjunto de datos. Este tipo de estrategias engloban principalmente las estrategias señaladas en (Kuhn, Max and Johnson 2013) y que se describen en las siguientes tablas (de Tabla 5.1 a Tabla 5.4): técnicas de integración, técnicas de limpieza, técnicas de transformación y técnicas de reducción de datos.

El uso de las diferentes técnicas de pre-procesamiento depende de la naturaleza de las medidas, del método de adquisición utilizado y del procesamiento final de las mismas. Por ejemplo, si se trabaja con datos provenientes de varias fuentes será necesario el uso de estrategias de integración mientras que las estrategias de reducción tendrán sentido cuando se trabaje con un gran volumen de datos.

Tabla 5.1. Estrategia de integración de datos

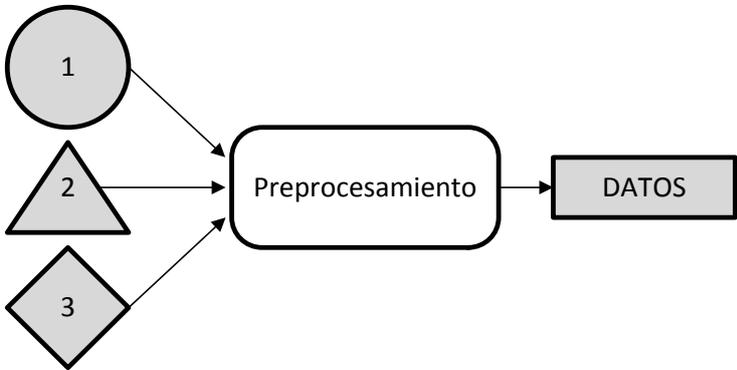
Estrategia	Problemas a resolver	Procedimientos
Integración de datos	<ul style="list-style-type: none"> - Diferentes fuentes de información - Diferentes formas de representación y codificación 	<ul style="list-style-type: none"> - Análisis de redundancia - Análisis de correlación
Descripción		
<p>Partiendo de datos provenientes de diferentes fuentes se realizan las transformaciones o modificaciones oportunas para poder almacenar los datos sin errores en un mismo repositorio</p> 		

Tabla 5.2. Estrategia de limpieza de datos

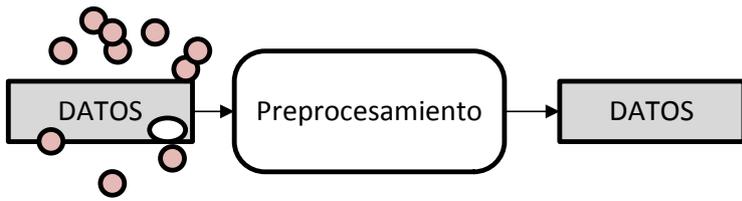
Estrategia	Problemas a resolver	Procedimientos
Limpieza de datos	<ul style="list-style-type: none"> - Ausencia de valores - Datos incoherentes - Ruido en los datos 	<ul style="list-style-type: none"> - Eliminación de tuplas - Auto-completado de datos - Filtrado
Descripción		
<p>Assumiendo un origen de datos ruidoso con datos ausentes y datos incoherentes, se tendrán a la salida los datos originales coherentes, completos y sin ruido</p> 		

Tabla 5.3. Estrategia de transformación de datos

Estrategia	Problemas a resolver	Procedimientos
Transformación de datos	<ul style="list-style-type: none"> - Datos no representativos - Datos difícilmente interpretables 	<ul style="list-style-type: none"> - Suavizado - Agregación - Normalización
Descripción		
<p>A partir de los datos originales, se construyen nuevos atributos o características que simplifican el procesamiento y que dotan al sistema de una mayor información que los datos originales</p>		
<pre> graph LR A((DATOS)) --> B[Preprocesamiento] B --> C[DATOS] </pre>		

Tabla 5.4. Estrategia de reducción de datos

	Problemas a resolver	Procedimientos
Reducción de datos	<ul style="list-style-type: none"> - Cantidad de datos elevada 	<ul style="list-style-type: none"> - Reducción dimensional - Comprensión de datos - Discretización
Descripción		
<p>Se disminuye el volumen de datos para trabajar con una representación reducida del conjunto original pero que produce los mismos (o casi iguales) resultados analíticos</p>		
<pre> graph LR A[DATOS] --> B[Preprocesamiento] B --> C[DATOS] </pre>		

Las técnicas de transformación son las estrategias más utilizadas en el preprocesamiento de la medida del comportamiento. El objetivo principal del uso de estas técnicas es la eliminación de datos incoherentes, la corrección de errores y la completitud de datos faltantes.

5.2.2.2. SEGMENTACIÓN

El objetivo de la etapa de segmentación es dividir la variable a estudiar en varias partes que permitan simplificar y mejorar su procesamiento. En el caso de variables continuas, como es la actividad física, se identifican en el estado del arte las tres técnicas principales de segmentación que se muestran en la Tabla 5.5.

Tabla 5.5. Principales técnicas de segmentación

Técnica	Ventajas	Inconvenientes
Enventanado temporal	Intuitivo, sencillo y cumple requisitos de tiempo real	Posible pérdida de patrones característicos
División temporal de la variable	<p>Algoritmos</p> <p>Existen diferentes tipos de ventana que pueden ser aplicados (Figura 5.5). Es necesario definir tamaño, forma y solapamiento</p> <p>Referencias</p> <p>(Parkka and Ermes 2006) (Fontecha and Hervás 2011) (Zhou, Chen, and Chung 2008) (Iglesias and Angelov 2010)</p>	
Técnica	Ventajas	Inconvenientes
Detección de regiones	Evita la separación de patrones característicos de la variables	No cumple requisitos de tiempo real (complejidad alta)
División de la variable según regiones identificadas.	<p>Algoritmos</p> <p>Destacan los algoritmos basados en umbralización, las técnicas de detección de patrones como la transformada de Hough y los algoritmos <i>top-down</i> y <i>bottom-up</i></p> <p>Referencias</p> <p>(Molleda et al. 2012) (Ye, Dobson, and McKeever 2012) (Thomas et al. 2010) (Medjahed and Istrate 2009)</p>	
Técnica	Ventajas	Inconvenientes
Detección de fronteras	La segmentación tiene lugar en base a los cambios detectados	Coste computacional alto y necesidad de calibración
División de la variable según bordes detectados	<p>Algoritmos</p> <p>Principalmente algoritmos de detección de bordes como el detector de Canny o las técnicas de análisis escala-espacio</p> <p>Referencias</p> <p>(Senthilkumaran and Rajesh 2009) (Heric and Zazula 2007)</p>	

La etapa de segmentación cobra gran importancia, por ejemplo, en el procesamiento de imágenes ya que permite segmentar la información en función de la aparición de un objeto o de las características de color o textura de la imagen (Peng, Zhang, and Zhang 2013), (Sharma and Aggarwal 2010), (Ilea and Whelan 2011). En un sistema de monitorización de personas mayores, esta etapa permite dividir la variable monitorizada en segmentos de menor tamaño para procesar la información de forma más eficiente y efectiva. En el diseño se deben tener en cuenta los requisitos de tiempo real necesarios en la monitorización, el tamaño de los datos que se pre-procesan y su velocidad de generación (Anuradha and Roy 2014).

De entre todas las técnicas de segmentación descritas, en el campo del comportamiento humano destacan los algoritmos de enventanado temporal. En este tipo de algoritmos es necesario establecer tres parámetros principales:

- *Tamaño de la ventana.* Es el número de muestras de la variable que formarán parte de cada segmento (ver Figura 5.4). Generalmente suele ser un número prefijado e igual para todos los segmentos aunque existen trabajos como (Keegan et al. 2013) o (Ikonomovska et al. 2011) que proponen un enventanado adaptativo que permita variar el número de muestras en función de las características de la variable.

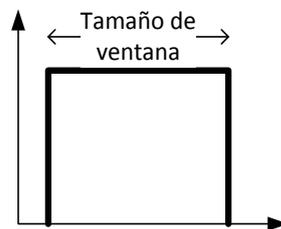


Figura 5.4. Tamaño de ventana

- *Forma de la ventana.* Aunque generalmente se trabaja con ventanas cuadradas, existen diferentes tipos de ventanas que pueden ser utilizados en la segmentación (Chakraborty and Kovvali 2013). La Figura 5.5 muestra los tres tipos principales de ventana utilizados en el estado del arte: ventana cuadrada, ventana triangular y ventana de Hamming. Mientras que la ventana cuadrada divide la variable en segmentos de tamaño constante sin variar los valores de las muestras originales, el enventanado triangular y el enventanado de Hamming realizan una transformación de la variable dejando la muestra central de la ventana en su estado original pero atenuando el resto siguiendo diferentes patrones.

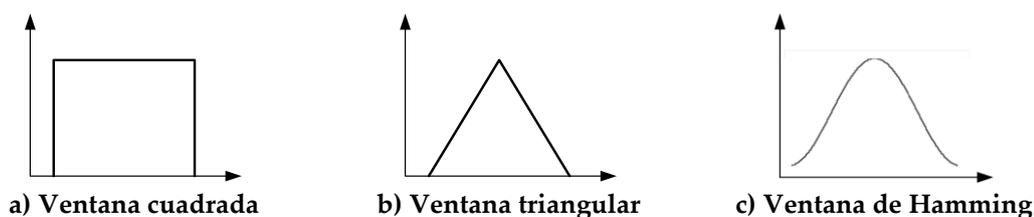


Figura 5.5. Tipos de ventanas que pueden ser utilizadas en la segmentación

- *Muestras de solapamiento.* La segmentación de una variable puede tener lugar sin solapamiento o con solapamiento. Un enventanado con solapamiento implica que parte de las muestras utilizadas en un segmento son reutilizadas en el siguiente segmento. La segmentación con solapamiento puede ser total, si no existen muestras que queden fuera de la siguiente ventana o parcial si solamente se solapan parte de las muestras (ver Figura 5.6.b). Un enventanado tiene lugar sin solapamiento (ver Figura 5.6.a) si las ventanas son consecutivas y no comparten ninguna muestra.

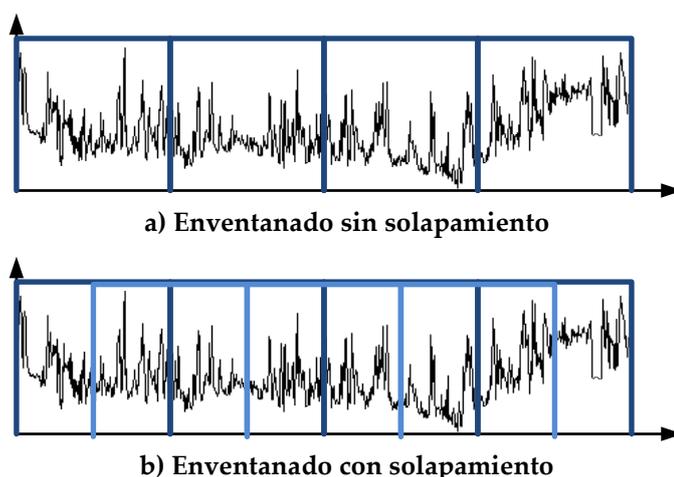


Figura 5.6. Enventanado con y sin solapamiento

A la hora de utilizar un enventanado temporal para realizar la segmentación de una variable es necesario seleccionar una forma de ventana que se adecúe a la aplicación y llegar a un compromiso entre el tamaño de la ventana y las muestras de solapamiento. Si se trabaja con un tamaño de ventana muy grande y sin solapamiento la segmentación de la variable será muy pequeña y los resultados en el procesamiento serán similares a trabajar con la variable total. Sin embargo, si se trabaja con un tamaño de ventana muy pequeño y solapamiento total, la complejidad del algoritmo de procesamiento aumentará y el objetivo principal de la segmentación no se conseguirá.

5.2.2.3. EXTRACCIÓN DE CARACTERÍSTICAS

Una vez que la variable es segmentada, es necesario extraer características de cada segmento que puedan ser utilizadas en el procesado y que sean de interés para el modelado (Avci and Bosch 2010). El tipo de características extraídas dependen del dominio de trabajo y del objetivo de la monitorización y pueden destacarse los tres tipos descritos en la Tabla 5.6: características temporales, características espectrales y características propias del dominio. Para cada tipo de características se muestran algunos ejemplos y referencias a su utilización en el estado del arte.

Las características temporales son las características más utilizadas para el modelado ya que su estudio permite realizar una correcta caracterización de la variable y presentan requisitos computacionales más bajos que las características espectrales (Figo et al. 2010) (Yang 2009) (Curone et al. 2010).

Tabla 5.6. Principales tipos de características

Tipo	Ejemplos	Referencias
Características temporales Características matemáticas y/o estadísticas de la variables	Media, moda, mediana, máximo, mínimo, varianza, desviación estándar, desviación media, rango de medida, coeficientes AR, relación entre variables, cálculo de pendientes	(Kwapisz, Weiss, and Moore 2011) (Lee et al. 2009) (Khan et al. 2010) (Long, Yin, and Aarts 2009) (Chen 2011)
Características espectrales Asumen la existencia de periodicidad en la variable y realizan su descomposición espectral	Componentes de la transformada de Fourier, energía espectral, entropía espectral, rango frecuencial, coeficientes de la transformada Wavelet...	(Zhang et al. 2012) (Hong et al. 2010) (Sharma and Gedeon 2012) (Doukas and Maglogiannis 2008) (Godfrey et al. 2008)
Características propias Dependientes del ámbito específico de trabajo	Número de pasos, pulso cardiaco, eventos detectados...	(Fleury et al. 2010) (Chung and Liu 2008) (Fletcher et al. 2010) (Khalil and Glal 2009)

5.2.3. MÉTODO DE PROCESADO DEL COMPORTAMIENTO

El objetivo de esta fase es la construcción de un modelo esperado de comportamiento que defina la rutina del usuario. Para ello existen tres tendencias principales en el estado del arte: métodos basados en modelado teórico, métodos basados en el reconocimiento de actividad y métodos basados en técnicas de aprendizaje automático (Avci and Bosch 2010).

Los métodos de procesado basados en modelado teórico se basan en la existencia de un modelo conocido o de conocimiento experto sobre el sistema. Según (Si et al. 2011) existen dos posibles alternativas para la construcción de un modelo sin utilizar los datos de medida:

- *Basadas en la existencia de un modelo.* Este tipo de modelado se puede realizar cuando los datos monitorizados se corresponden con algún fenómeno físico conocido que se rige por unas determinadas leyes. Este tipo de técnicas son difícilmente aplicables ya que la mayoría de los modelos existentes no se ajustan correctamente a los sistemas reales. En el campo del comportamiento humano, dentro de este tipo de modelos destacan el filtro de Kalman (Thobbi, Gu, and Sheng 2011) (Yuan and Chen 2012) y las teorías de comportamiento existentes en el campo de la psicología (Ajzen 1991) (Armitage and Conner 2001).
- *Basadas en conocimiento experto.* Se basan en la existencia y el uso de conocimiento experto acerca del sistema que se monitoriza. Este tipo de procesado suele utilizarse cuando no existen datos históricos del sistema que se modela y se basan en asunciones sobre el mismo. El modelo construido suele presentar bastante grado de error y su diseño e implementación son muy complejos. Su uso tiene sentido cuando se trabaja con medidas subjetivas o difícilmente medibles (Ellingsen et al. 2014) (Knudson 2013)

Tal y como se muestra en la Figura 5.7, en este tipo de métodos, el modelado tiene lugar sin tener en cuenta los datos de medida. Las entradas del método son los parámetros que definen el modelo teórico o el conocimiento experto que se tiene sobre el mismo. En general, este tipo de técnicas son una posible opción cuando se trabaja en un entorno sencillo y estático. Sin embargo, siempre suponen un alto coste en su implementación al ser necesario introducir manualmente reglas o ajustarse a modelos que definan el comportamiento (Xiang and Gong 2008).

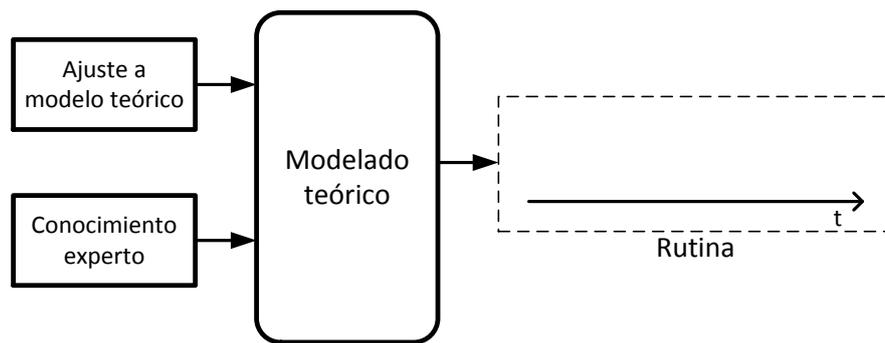


Figura 5.7. Método de procesamiento basado en un modelado teórico

En los métodos basados en el reconocimiento de actividad, las características extraídas en la etapa de acondicionamiento son utilizadas para determinar la actividad realizada por el usuario. El método de procesamiento determina la rutina identificando el instante temporal de comienzo y el instante temporal de final de cada una de las actividades esperadas para el usuario (Aggarwal and Ryoo 2011) (Biswas et al. 2010). Este tipo de técnicas son muy útiles y presentan una gran efectividad cuando existen actividades muy bien definidas y limitadas (Atallah and Yang 2009) (Chen and Khalil 2011). El modelo de comportamiento generado predecirá para cada instante temporal la actividad realizada por el usuario. Para ello, es necesaria una etapa exhaustiva de entrenamiento y etiquetado definiendo con antelación las posibles actividades susceptibles de ser detectadas. La descripción general de este tipo de procesamiento se muestra en la Figura 5.8.

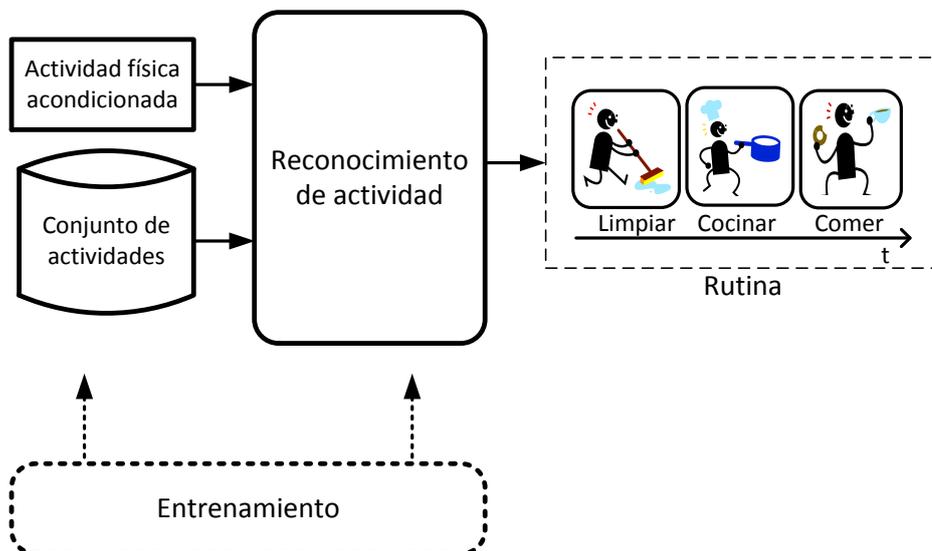


Figura 5.8. Método de procesamiento basado en reconocimiento de actividad

Tal y como muestra la Figura 5.9.a, la mayoría de los métodos para el reconocimiento de actividad asumen que cuando una actividad termina, comienza otra y así sucesivamente. En la realidad las actividades tienen lugar como se describe en la Figura 5.9.b y Figura 5.9.c. Una actividad puede tener lugar en paralelo con otra (Figura 5.9.c) o una actividad puede abandonarse momentáneamente para realizar otra (Figura 5.9.b).

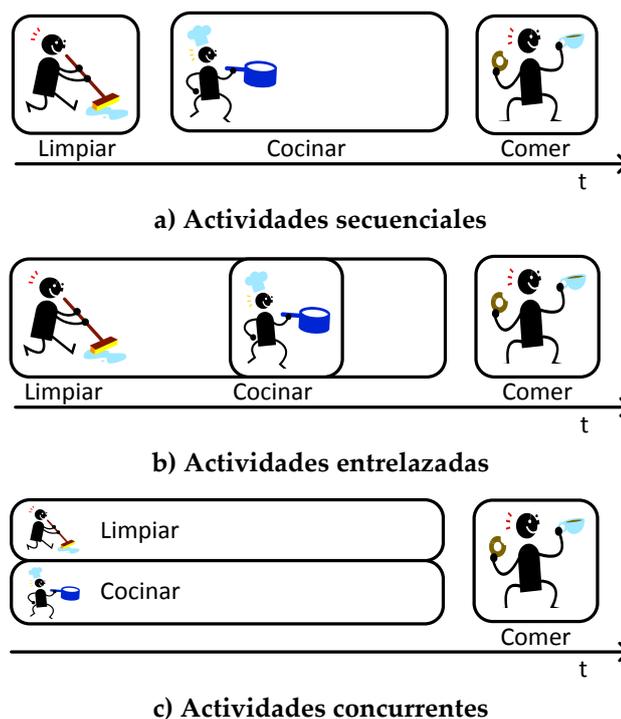


Figura 5.9. Escenarios de un sistema de reconocimiento de actividad

El reconocimiento de actividad en entornos reales presenta la detección y manejo de actividades concurrentes como uno de los principales retos a resolver tal y como se señala en (Lara and Labrador 2013) o (Kim, Helal, and Cook 2010), entre otros. Así, no se trata de una opción adecuada para ser utilizada en la construcción de un modelo de comportamiento ya que el error que conlleva el reconocimiento de actividad supondrá una importante fuente de error en el modelado. A su vez, ha de tenerse en cuenta la necesidad de partir de un conjunto cerrado de posibles actividades y las etapas de entrenamiento y etiquetado necesarias para su funcionamiento.

Partiendo del nivel de actividad física como la medida utilizada para la caracterización del comportamiento, las técnicas de aprendizaje automático son señaladas como una de las opciones más adecuadas para la identificación de rutinas (Jakkula and Cook 2008). Dentro de las técnicas de aprendizaje automático, destacan las técnicas de regresión.

Las técnicas de regresión asumen la existencia de un modelo subyacente en el sistema y su objetivo es la obtención de un conjunto de parámetros que lo definan y que permitan predecir los valores futuros de una serie de datos (Esling and Agon 2012). La Figura 5.10 muestra el esquema general de la aplicación de este tipo de técnicas para la identificación de rutinas.

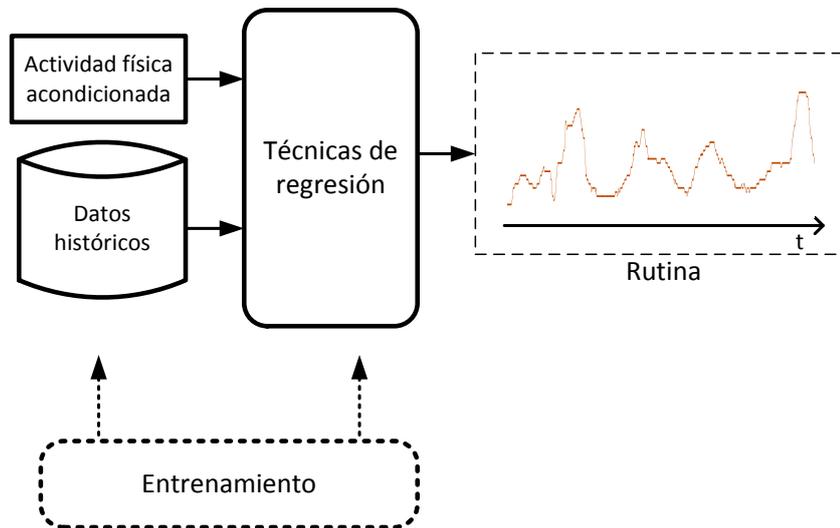


Figura 5.10. Método de procesamiento basado en técnicas de regresión

Siguiendo el esquema de la Figura 5.10, las técnicas de regresión parten de la actividad física medida y de la actividad física histórica para la identificación de un modelo de rutina que permita predecir para cada instante el nivel de actividad esperado. Dependiendo de la técnica de regresión seleccionada, puede ser requerida una etapa previa de entrenamiento.

Dentro de las técnicas de regresión destacan principalmente los métodos bayesianos, las redes neuronales, las cadenas ocultas de Markov y los métodos estadísticos (Kayacan, Ulutas, and Kaynak 2010). Todos estos tipos de técnicas se describen en la Tabla 5.7. Para cada una de ellas se presentan sus ventajas, inconvenientes y referencias a trabajos del estado del arte en los que se utilizan. De entre todas las técnicas mencionadas, los métodos estadísticos destacan dentro del modelado del comportamiento humano y son los que mejor se adaptan a los requisitos de este trabajo.

Tabla 5.7. Principales técnicas de regresión para el modelado

Técnica	Ventajas	Inconvenientes
Métodos bayesianos Los métodos bayesianos modelan la probabilidad de que se produzca una determinada salida dada una entrada al sistema	- La salida del sistema recibe un valor relacionado con la probabilidad de que sea cierta	- Puede verse sesgado por el conjunto de usuarios y/o observaciones utilizadas
	Referencias (Gupta, Kembhavi, and Davis 2009) (Lu and Fu 2009) (Xiang and Gong 2008)	
Técnica	Ventajas	Inconvenientes
Redes neuronales Sistema de aprendizaje y procesamiento automático basado en la existencia de unidades de proceso muy simples trabajando en paralelo	- Robustez al ruido en las entradas - Diferentes tipos : multicapa, jerárquicas	- Su correcto funcionamiento reside en la existencia de un conjunto cerrado y estudiado de posibles entradas
	Referencias (Trost and Wong 2012) (Staudenmayer et al. 2009) (Lin et al. 2012)	
Técnica	Ventajas	Inconvenientes
Cadenas ocultas de Markov Modelado estadístico en el que se asume que el proceso a modelar es un proceso de Markov de parámetros desconocidos	- Se añade información temporal modelando el paso de una actividad a otra	- Requiere entrenamiento - Se asume que un estado solo depende del anterior - El tiempo que se permanece en un estado no se utiliza para modelar la transición
	Referencias (Zhu and Sheng 2011) (Helaoui, Niepert, and Stuckenschmidt 2011)	
Técnica	Ventajas	Inconvenientes
Métodos estadísticos Se utilizan métodos estadísticos para generar un modelo que permita predecir los valores futuros	Automático sin necesidad de una gran etapa de entrenamiento	Se basa en asunciones estadísticas acerca de la naturaleza de los datos
	Referencias (Virone 2009) (Chen et al. 2013) (Vandewiele and Motamed 2011) (Mahmoud 2013)	

En (Chandola et al. 2009) se identifican dos aproximaciones principales dentro de los métodos estadísticos para el modelado:

- *Métodos paramétricos.* Asumen la existencia de un modelo paramétrico que define la distribución estadística que siguen los datos y que puede representarse como $f(x, \theta)$. Los parámetros que definen el modelo se definen como θ y f representa la función de densidad de probabilidad para una observación x .
- *Métodos no paramétricos.* No se realiza el ajuste de los datos a ninguna distribución conocida y el modelo obtenido se construye a partir de los propios datos medidos. Dentro de estos métodos se encuentra la regresión lineal y el cálculo de parámetros estadísticos.

La principal diferencia entre los métodos paramétricos y no paramétricos está en la asunción de que los datos pertenecen o no a una determinada distribución teórica. El uso de métodos paramétricos permite utilizar técnicas estadísticas robustas y ampliamente aceptadas pero es necesario definir a priori la distribución que siguen los datos. Esto supone un importante inconveniente, sobre todo cuando se trabaja con volúmenes grandes de datos o con fuentes variables.

Cuando la distribución de datos estudiada se ajusta a una distribución de datos conocida, el modelo de comportamiento es definido a través de los parámetros θ que caracterizan a la distribución. Este tipo de ajuste es de gran utilidad ya que permite el uso de métodos estadísticos y cálculos conocidos tanto para su construcción como para su posterior análisis. Una aproximación común en el estado del arte es el ajuste de los datos a una distribución normal de parámetros μ (media) y σ (desviación estándar). Este tipo de métodos destacan principalmente en el ámbito industrial para el control de la calidad de procesos pero son también utilizados en el ámbito de la salud (Gokalp and Clarke 2013). Para realizar este ajuste, en ocasiones es necesario realizar una transformación de los datos originales para conseguir llegar a un conjunto de datos que sí se ajusten correctamente.

La principal ventaja de los métodos no paramétricos es que no necesitan realizar un ajuste de los datos a una distribución conocida y se basan en los propios datos de medida para generar el modelo. En este tipo de métodos, un cambio en la fuente de datos o en la distribución de los mismos no requiere de un reentrenamiento o de una reconfiguración del sistema.

5.3. DISEÑO DEL MÉTODO DE MODELADO

5.3.1. MODELO ESPERADO (RF-6)

En la búsqueda de un método de modelado automático y adaptativo (requisitos identificados en el Capítulo 2) se debe buscar un método que identifique la rutina característica de cada usuario y que pueda ser recalculado de forma automática para adaptarse temporalmente a los posibles cambios. Así, la técnica utilizada deberá basarse en datos históricos para la generación de la variable de salida.

Dependiendo de la técnica utilizada para el modelado, la rutina será definida a través del conjunto de actividades que realiza el usuario (técnicas basadas en el reconocimiento de actividad) o a través del nivel de actividad realizado (técnicas de aprendizaje automático). Las técnicas basadas en el reconocimiento de actividad presentan dos limitaciones principales que hacen que no cumplan los requisitos buscados en este trabajo. Este tipo de técnicas, tal y como ya se mencionó en la sección 5.2.3, no son capaces de definir adecuadamente la naturaleza del comportamiento humano. Además, se basan generalmente en técnicas de clasificación y requieren una etapa previa de entrenamiento con etiquetado que permita extraer conocimiento para la clasificación y reconocimiento de actividad. Esto hace que se conviertan en técnicas difícilmente escalables y no adaptables al no ser capaces de adaptarse a los posibles cambios que puedan tener lugar en el sistema.

Las técnicas de aprendizaje automático permiten crear modelos para la predicción del comportamiento futuro de un sistema. Este tipo de técnicas son recomendadas para el modelado del comportamiento humano en algunos trabajos como (Butte, Ekelund, and Westerterp 2012), (Ye et al. 2012) o (Xiang and Gong 2008). Dentro de este tipo de técnicas destacan los métodos estadísticos. Este tipo de métodos, al contrario que las técnicas de reconocimiento de actividad, pueden ser ejecutados de forma automática y adaptativa.

El modelo esperado, será una variable que pueda ser interpretada como la rutina del usuario y que permita conocer para cada instante temporal cuál es el valor de actividad física esperado. Así, el modelo esperado será una variable similar a la representada en la Figura 5.11 fácilmente interpretable y en las mismas unidades que la variable de medida.

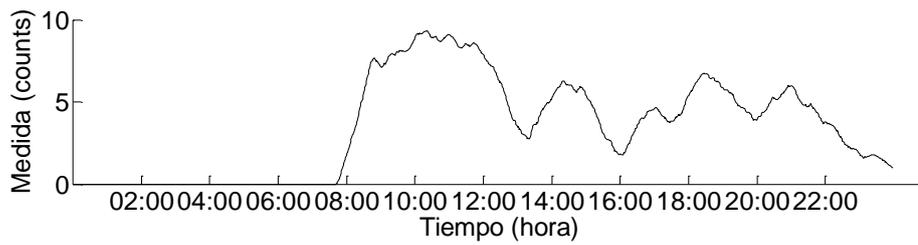


Figura 5.11. Ejemplo de modelo de comportamiento esperado (rutina)

5.3.2. MÉTODO DE ACONDICIONAMIENTO (RF-4)

5.3.2.1. INTRODUCCIÓN

La entrada al método de acondicionamiento es la actividad física obtenida mediante el método de medida descrito en el Capítulo 4 y la salida será la variable de entrada al método de procesado (descrito en la sección 5.3.3). Tanto el método de acondicionamiento como el método de procesado trabajan con días completos de datos como unidad mínima de procesamiento. Esta consideración está ampliamente extendida y está relacionada con el ritmo circadiano del comportamiento humano (Gokalp and Clarke 2013) (Fouquet and Franco 2010).

El método de acondicionamiento se divide en una fase de pre-procesamiento (descrita en la sección 5.2.2.1), una fase de segmentación (descrita en la sección 5.2.2.2) y una fase de extracción de características (descrita en la sección 5.2.2.3).

5.3.2.2. MÉTODO DE PRE-PROCESAMIENTO

Esta fase del método de acondicionamiento permite resolver problemas ocasionados por datos ausentes, datos incoherentes o datos con ruido. Tal y como se describió en la sección 5.2.2, dentro de esta fase destacan las técnicas de integración, limpieza, transformación y reducción de datos.

Las técnicas de integración tienen sentido cuando se trabaja con datos provenientes de diferentes fuentes. En el método de medida diseñado en el Capítulo 4 no se tiene en cuenta la posibilidad de trabajar con más de una fuente de información, por lo que este tipo de técnicas no son aplicables a este trabajo. Si se trabajara simultáneamente con más de un dispositivo medidor de actividad, sí podría ser necesario implementar técnicas de integración que garantizaran principalmente la sincronización de datos entre fuentes de información. Dentro de las técnicas de transformación aplicadas a la actividad física, destaca la técnica de normalización que permitiría realizar la comparación de medidas provenientes de diferentes dispositivos. Sin embargo, puesto

que no se baraja este tipo de posibilidad, este tipo de técnicas tampoco son consideradas en este trabajo.

El método de medida de la actividad física (cálculo de la variable JIM) genera un dato cada minuto, lo que supone la generación de 1440 valores a lo largo de un día (unidad de procesamiento). El tamaño de estos datos dependerá de la resolución del acelerómetro integrado en el dispositivo vestible. Asumiendo un tamaño de 2 bytes por dato, el volumen de datos generado es de 2880 bytes. Teniendo en cuenta que será necesario incorporar información temporal en las muestras, se podría concluir que el método de medida de la actividad física genera diariamente menos de 3 MB de datos (ver el cálculo representado en la Ecuación 5.1). Puesto que se trata de un volumen de datos bajo, no es necesario implementar técnicas de reducción de datos en el método de acondicionamiento.

$$1 \frac{\text{muestra}}{\text{minuto}} \times 60 \frac{\text{minutos}}{\text{hora}} \times 24 \frac{\text{horas}}{\text{día}} \times 2 \frac{\text{bytes}}{\text{muestra}} = 2880 \text{ bytes} < 3 \text{ MB}$$

Ecuación 5.1

Intrínsecamente, el uso de dispositivos electrónicos en la medida conlleva la inclusión de ruido y la posibilidad de aparición de muestras incoherentes. A su vez, en ocasiones pueden tener lugar fallos en el sistema que conlleven la ausencia de datos y que deberán ser tratados mediante alguna técnica de limpieza de datos. La ausencia de datos puede deberse a diferentes factores relacionados con el sistema de adquisición (tiempo de vida, fallos) o con el usuario monitorizado (rechazo a ser monitorizado). Resolver este tipo de situaciones es uno de los principales retos de un sistema de monitorización según (Masse et al. 2005) o (Fouquet and Franco 2010).

Para resolver esta limitación, en este trabajo se ha diseñado una estrategia de limpieza de datos que podría ser aplicada de forma general sobre cualquier sistema de monitorización. La estrategia propuesta se basa en el uso combinado de una técnica de eliminación de datos redundantes y una técnica de detección de ausencia de datos. Ambas técnicas son aplicadas sobre la medida durante un día completo de monitorización:

- *Técnica de eliminación de datos redundantes.* Se analiza la existencia de más de un instante temporal con un valor de medida. Este tipo de redundancia tiene lugar debido a fallos en el método de medida y en la sincronización temporal del dispositivo utilizado. En el caso de que exista un instante temporal con más de un dato de medida, se eliminará.

- *Técnica de detección de datos ausentes.* Se analiza la existencia de instantes temporales sin datos de medida. En función de la cantidad de datos faltantes se propone el uso de una técnica de eliminación o una técnica de auto-completado para su resolución:
 - o *Técnica de eliminación de datos.* Si en un día no existen datos suficientes para su procesamiento, ese día será eliminado completamente.
 - o *Técnica de auto-completado de datos.* Las muestras faltantes serán estimadas para poseer la medida de la actividad física durante todos los instantes temporales de monitorización.

En la implementación de una técnica para la detección de datos ausentes es necesario definir una mínima cantidad de datos que permita decidir la eliminación total o el auto-completado de los datos existentes. Esa mínima cantidad de datos (N_{min}) definirá las características de un día válido para la monitorización. En (Catellier et al. 2005) se establece la necesidad de tener datos de al menos un 70% del día para optar por la no eliminación del mismo. Siguiendo esta idea, un día será descartado en el pre-procesamiento cuando haya más de 7 horas sin datos (> 420 muestras sin datos). En otro caso, se optará por aplicar una técnica de auto-completado que permita generar datos de información para el día completo.

Dentro de las técnicas de auto-completado en la medida de la actividad física destaca el uso de técnicas de interpolación y de técnicas basadas en el uso de la media de las observaciones (Warren et al. 2010). En otros trabajos como (Catellier et al. 2005), (Scarmeas et al. 2009) o (Hamm et al. 2013) se describen otro tipo de técnicas que pueden ser utilizadas para el auto-completado de datos. Por ejemplo, en (Catellier et al. 2005) se opta por el uso de estrategias basadas en técnicas probabilísticas.

En este trabajo se propone el uso de la media de las observaciones como técnica para el auto-completado de datos faltantes. Considerando el día como unidad de monitorización, los datos ausentes serán completados con la media de los datos observados durante los últimos N_r días para ese instante temporal. La estrategia propuesta en este trabajo para el auto-completado simplifica el proceso y ha demostrado su validez y buenos resultados cuando los datos ausentes se encuentran en torno al 10% (Catellier et al. 2005). Esta técnica evita los posibles errores con los que se lidia cuando se trabaja con días con pocos datos y a su vez permite reducir el sesgo introducido cuando se eliminan días con pocos datos ausentes.

La Figura 5.12 muestra el diagrama de bloques del método de pre-procesamiento propuesto en este trabajo y descrito en esta sección. El método propuesto se basa en la combinación de una estrategia de eliminación de datos redundantes junto con una estrategia combinada de eliminación y auto-completado de datos. El número de días utilizados para la obtención de la media es fijada en $N_r = 5$ de acuerdo a las asunciones de rutina llevadas a cabo en algunos trabajos como (Heil et al. 2012) o (Trost, Mciver, and Pate 2005).

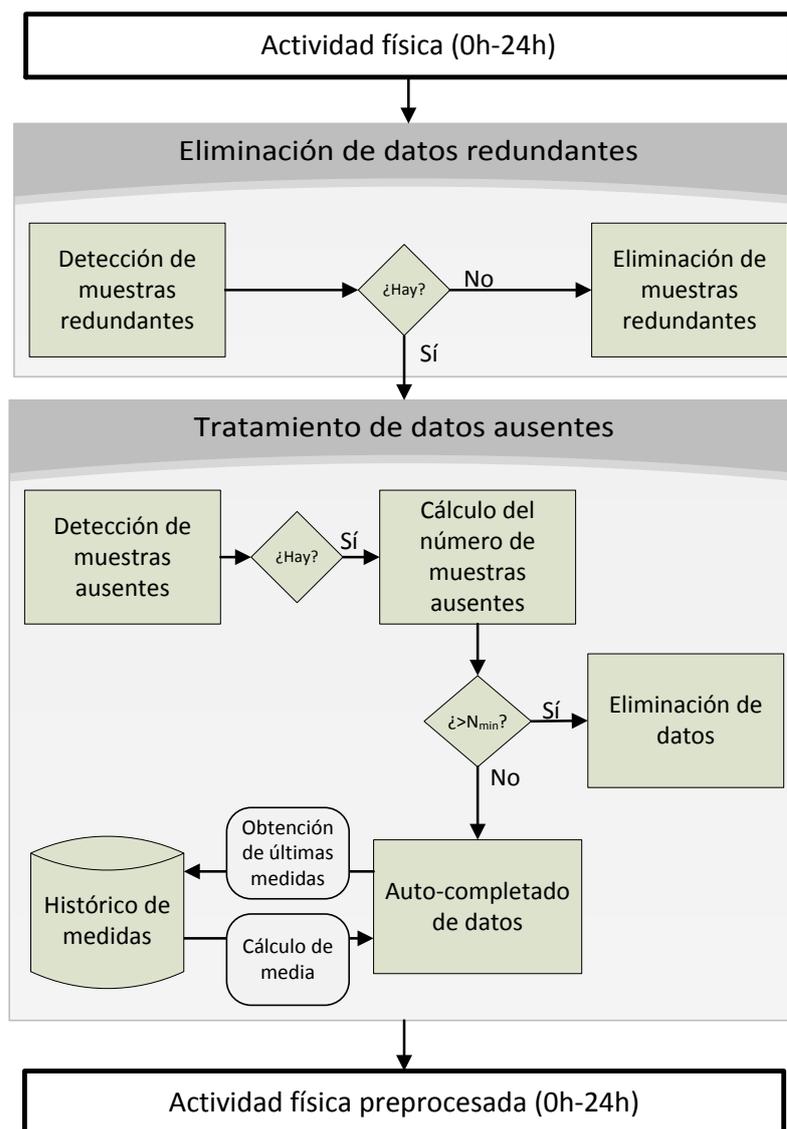


Figura 5.12. Método de pre-procesamiento propuesto: técnica de limpieza de datos

La implementación de la estrategia propuesta para la eliminación de datos redundantes (ver la Figura 5.13) se basa en la comprobación de que solo existe una muestra de actividad para cada uno de los instantes de monitorización. Para ello, el método *preprocesarDatosRedundantes* pide la actividad física medida (*obtenerActividadFísica*) para cada hora (*h*) y minuto (*m*) del día (*d*) sobre el que se desea aplicar la estrategia. En el caso de recibir más de una muestra de actividad se invocará el método *eliminarActividadFísica* con el objetivo de eliminar las muestras redundantes existentes en la hora *h* y minuto *m*.

```

Funcion preprocesarDatosRedundantes(Día d):
  For Integer h from 0 to 23:
    For Integer m from 0 to 59:
      Double[] Actividad = obtenerActividadFísica(d, h, m) ;
      If (length(Actividad) > 1):
        // Hay muestras redundantes
        For Integer i from 1 to length(Actividad)-1:
          eliminarActividadFísica(d, h, m);
        End for;
      End if;
    End for;
  End for;
End function;

```

Figura 5.13. Pseudocódigo de la estrategia de eliminación de datos redundantes

El método *preprocesarDatosAusentes*, cuyo pseudocódigo se muestra en la Figura 5.14, representa la estrategia propuesta para detección de datos ausentes y su eliminación o auto-completado. Dentro del día *d* sobre el que se desea realizar la detección, se pide la actividad física medida (*obtenerActividadFísica*) para cada hora (*h*) y minuto (*m*). En la matriz *DatosAusentes* se almacenarán todos los instantes [*h*, *m*] para los cuales no hay datos de actividad.

En función del número de datos existentes y su relación con el mínimo (N_{min}) se tomará la determinación de aplicar una estrategia de eliminación de datos (*eliminarActividadFísica*) o una estrategia de auto-completado (*completarActividadFísica*). Esta última consiste en el cálculo de la media de la actividad física medida durante los N_r días anteriores al día que se está analizando en los mismos instantes [*h*, *m*] para los cuales no hay datos. Los nuevos valores calculados serán introducidos en la base de datos de medidas a través del método *modificarActividadFísica*.

```

Function preprocesarDatosAusentes(Día d, Integer  $N_{\min} = 1020$ , Integer  $N_r = 5$ ):
    Integer NumDatosAusentes = 0;
    Integer[] DatosAusentes;
    // Se pide la actividad física medida para cada hora h y minuto m del día d
    For Integer h from 0 to 23:
        For Integer m from 0 to 59:
            Double[] Actividad = obtenerActividadFísica(d, h, m) ;
            If (length(Actividad) == 0):
                // No hay muestras en la hora h y el minuto m
                DatosAusentes[NumDatosAusentes] = [h, m]
                NumDatosAusentes++;
            End if;
        End for;
    End for;
    If (NumDatosAusentes  $\geq 1440 - N_{\min}$ ):
        // Estrategia de eliminación de datos
        eliminarActividadFísica(d);
    Else:
        // Estrategia de auto-completado de datos
        completarActividadFísica(d, DatosAusentes,  $N_r$ );
    End if;
End function;

Function completarActividadFísica(Día d, Integer DatosAusentes[], Integer  $N_r = 5$ ):
    // Se recorre el conjunto de datos ausentes
    For Integer i from 1 to length(DatosAusentes):
        Integer h = DatosAusentes[i][1];
        Integer m = DatosAusentes[i][2];
        Double SumaActividadFísica = 0;
        Double MediaActividadFísica = 0;
        // Cálculo de la media de las medidas en los  $N_r$  días anteriores
        // para la hora h y el minuto m
        For Integer j from 1 to  $N_r$ :
            SumaActividadFísica += obtenerActividadFísica(d-j, h, m);
        End for;
        MediaActividadFísica = SumaActividadFísica/ $N_r$ ;
        // Se inserta nueva actividad física para el día d, la hora h y el minuto m
        modificarActividadFísica(d, h, m, MediaActividadFísica);
    End for;
End function;

```

Figura 5.14. Pseudocódigo de la estrategia de detección de datos ausentes

5.3.2.3. MÉTODO DE SEGMENTACIÓN

El objetivo general de esta etapa del método de acondicionamiento es dividir la medida original en segmentos de menor tamaño cuyo procesamiento pueda ser más útil y sencillo. A su vez, el objetivo del método de segmentación en este trabajo es transformar cada muestra de la variable en un conjunto de muestras que caractericen su vecindad. La finalidad es que cada muestra en el procesado represente el valor tomado en la vecindad y no solo en el instante temporal estudiado.

En la sección 5.2.2.2 se describieron las principales aproximaciones existentes en el estado del arte para la segmentación: inventariado temporal, detección de regiones y detección de fronteras (ver Tabla 5.5). En la sección E.1 del Apéndice E se muestran varios ejemplos de aplicación de los dos últimos tipos de técnicas mencionados.

En general, las técnicas de detección de regiones y de detección de fronteras son interesantes cuando se va a realizar un procesamiento diferente de la variable en función de su valor o de sus características. Este tipo de técnicas no son válidas para su aplicación en este trabajo por dos razones principales. En primer lugar, porque no se desea que el modelado de la actividad física varíe en función de sus características y en segundo lugar por la ausencia de conocimiento acerca de la forma que toma la variable y que hace que la implementación de un método de detección de patrones o fronteras no sea adecuado. La aparición de nuevos patrones o nuevos comportamientos en la actividad física medida implicaría el rediseño del método de segmentación.

En este trabajo, se opta por el diseño de una técnica de inventariado temporal cuyos parámetros se definen en la Tabla 5.8. El método de segmentación propuesto destaca principalmente por su sencillez, su fácil implementación y la no necesidad de conocimiento experto o asunciones acerca de los patrones o naturaleza de la variable medida. Se utilizará solapamiento total en la segmentación para garantizar que ningún patrón característico de la variable quede dividido en segmentos diferentes para su estudio. Así, se resuelve una de las principales limitaciones de los métodos de inventariado temporal.

Tabla 5.8. Parámetros del método de segmentación propuesto

Parámetro	Valor
Tamaño de la ventana	$N_m = 60$ muestras
Forma de la ventana	Cuadrada
Muestras de solapamiento	Solapamiento total

La Tabla 5.8 recoge los parámetros que definen el método de segmentación. Al tratarse de un método de segmentación con solapamiento total, el número de segmentos a la salida siempre será igual al número de muestras de entrada. La medida de la actividad física será dividida en segmentos de N_m muestras. Para cada muestra n de actividad se construirá el segmento $(n-N_m/2, n+N_m/2)$ por lo que a la salida del método siempre se tendrá un número de segmentos igual al número de muestras que se pre-procesan (1440 muestras, 24 horas). El valor N_m deberá ser seleccionado en función de la resolución deseada en el algoritmo de procesamiento. Se propone, un valor de 60 muestras dividiendo así la actividad en segmentos de 1 hora. Este valor para la segmentación es común en varios trabajos como (Gil et al. 2013) (Intille et al. 2011) (Mathie et al. 2004) o (Leon et al. 2010) y se ha validado, por ejemplo, en (Trost and Wong 2012). Como el procesamiento tiene lugar a nivel de día, las muestras iniciales y finales de la variable serán completadas con $n-N_m/2$ valores nulos para garantizar que todos los segmentos poseen el mismo tamaño.

En la última parte de la sección E.1 del Apéndice E se muestra la aplicación de diferentes tipos de formas de ventana sobre una variable de ejemplo de medida de la actividad. El objetivo de la segmentación en este trabajo es obtener para cada muestra, una representación de su vecindad. Por esta razón, se opta por una segmentación de ventana cuadrada.

La Figura 5.15 muestra el diagrama de bloques del método de segmentación propuesto en este trabajo. La actividad física pre-procesada será la entrada a un método de inventariado temporal que utilizará una ventana cuadrada de N_m muestras para la segmentación con solapamiento total de la variable de entrada. La salida del método de segmentación serán 1440 segmentos de tamaño N_m muestras.

La Figura 5.16 muestra el pseudocódigo que define el método de segmentación propuesto. El método *segmentarActividadFísica* recibe el día sobre el que se desea realizar la segmentación (d) junto con el número de muestra sobre el que se desea trabajar (n) y el tamaño deseado en la segmentación (N_m). En base a ese tamaño se pedirán las muestras de actividad que deben incluirse en el segmento y que será la actividad medida entre las muestras $n-N_m/2$ y $n+N_m/2$. En el caso de que no existan muestras (comienzo o final del día), el segmento será completado con valores nulos. Este método (*segmentarActividadFísica*) es invocado para cada una de las muestras n de la actividad física medida en el día d .

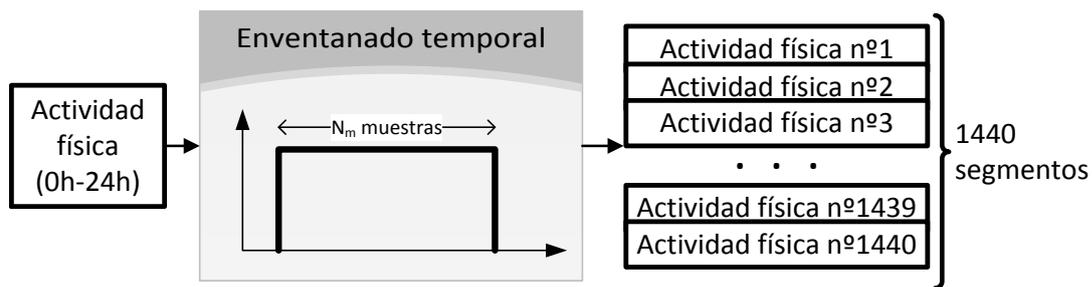


Figura 5.15. Método de segmentación propuesto: inventariado temporal

```

Function segmentarActividadFísica(Día d, Integer n, Integer Nm = 60):
    Integer Índice = 0;
    Double[Nm] SegmentoActividad;
    // Se obtiene segmento de actividad de tamaño Nm para la muestra n
    For Integer i from n-Nm/2 to n+Nm/2:
        If ( i < 1):
            // Si está en el inicio del día y no hay muestras, se completa con 0
            SegmentoActividad[Índice] = 0;
            Índice++;
        ElseIf ( i > 1440):
            // Si está en el final del día y no hay muestras, se completa con 0
            SegmentoActividad[Índice] = 0;
            Índice++;
        Else:
            // Se completa el segmento con los valores medidos
            // en el día d y en la muestra i
            SegmentoActividad[Índice] = obtenerActividadFísica(d, i);
            Índice++;
        End if;
    End for;
    Return SegmentoActividad;
End function;
    
```

Figura 5.16. Pseudocódigo del método de segmentación propuesto

5.3.2.4. MÉTODO DE EXTRACCIÓN DE CARACTERÍSTICAS

La fase de extracción de características trabaja sobre cada uno de los segmentos extraídos en la fase anterior obteniendo características útiles para el modelado y que permitan simplificarlo y simultáneamente aumentar su eficiencia y su eficacia.

Las características temporales y las características espectrales son las dos aproximaciones principales que destacan en el estado del arte. Ambas han mostrado su eficiencia y su efectividad en numerosos trabajos y en diferentes ámbitos tanto del procesamiento del comportamiento humano como en otras áreas científicas. Las

características temporales destacan sobre las características espectrales al presentar una menor complejidad computacional alcanzando los mismos resultados.

Dentro de las técnicas de extracción de características temporales, en este trabajo se opta por el uso de la media (ver Ecuación 5.2). El uso de la media es una técnica ampliamente utilizada como valor característico de una serie de datos cuantitativos partiendo del principio de la esperanza matemática. El principal inconveniente del uso de la media como valor característico de una variable es que su cálculo es muy sensible a los valores extremos presentes en la variable: la presencia de valores muy altos aumentan su valor y la presencia de valores muy bajos lo disminuyen considerablemente.

$$Media = \frac{\sum_{i=1}^N Muestra_i}{N}$$

Ecuación 5.2

El método propuesto en este trabajo se basa en el cálculo de la media sobre cada uno de los segmentos de N_m muestras extraídos en la fase anterior del acondicionamiento de la variable. Al estar trabajando con un inventariado temporal de solapamiento total, se consigue resolver la principal limitación del uso de la media como característica. La aparición de valores altos o bajos se verá amortiguada a lo largo del tiempo y la sensibilidad del cálculo matemático será insignificante.

El diagrama de bloques de la Figura 5.17 muestra el método de extracción de características propuesto en esta sección. Para cada uno de los 1440 segmentos extraídos en la fase de segmentación, se realizará el cálculo de la media siguiendo la expresión de la Ecuación 5.2. Así, a la salida de esta fase se tendrán 1440 valores.

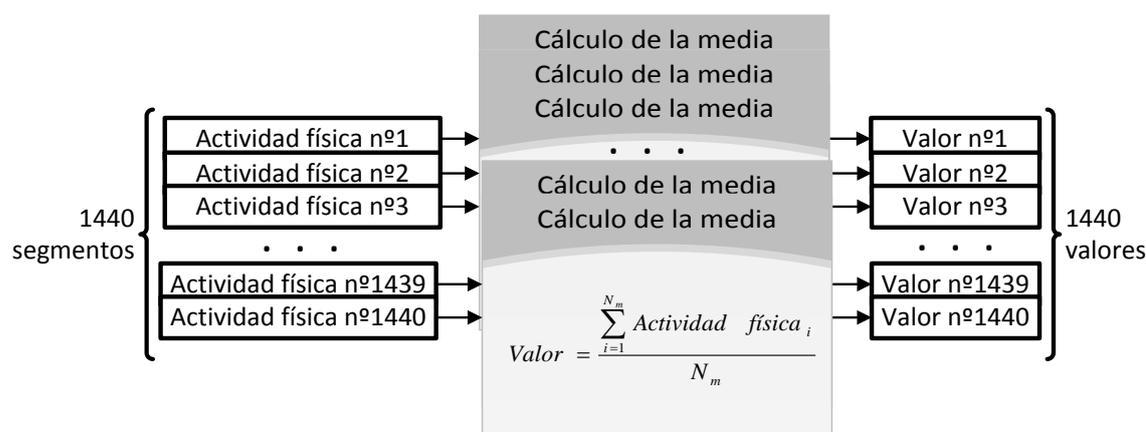


Figura 5.17. Método de extracción de características propuesto: cálculo de la media

El pseudocódigo de la Figura 5.18 representa la implementación del método de extracción de características propuesto. Dado un segmento de actividad (*SegmentoActividad*), el método *extraerCaracterísticas* devuelve la media de la actividad física medida. Este método debe ser invocado para cada uno de los segmentos identificados en la etapa de segmentación.

```
Function extraerCaracterísticas(Double[] SegmentoActividad):  
    Double SumaActividad = 0;  
    // Se calcula la media de las medidas contenidas en el segmento  
    For Integer i from 1 to length(SegmentoActividad):  
        SumaActividad += SegmentoActividad[i];  
    End for;  
    Return SumaActividad/length(SegmentoActividad);  
End function;
```

Figura 5.18. Pseudocódigo del método de extracción de características propuesto

5.3.2.5. RESUMEN

La Figura 5.19 muestra el diagrama de bloques del método propuesto para el acondicionamiento de la actividad física y la Figura 5.20 el pseudocódigo correspondiente.

El método *calcularActividadFísicaAcondicionada* devuelve para cada día *d* la actividad física acondicionada tras aplicar los métodos propuestos para el pre-procesamiento. Estos métodos se basan en la aplicación de una estrategia de eliminación de datos redundantes junto con el tratamiento de datos ausentes (combinación secuencial del método *preprocesarDatosRedundantes* y del método *preprocesarDatosAusentes*), segmentación (*segmentarActividadFísica*) y extracción de características (*extraerCaracterísticas*).

El método de pre-procesamiento propuesto realiza un análisis de la actividad física eliminando datos redundantes y detectando la ausencia de datos. En función de la cantidad de datos ausentes se procederá a su auto-completado utilizando información histórica o a su eliminación. El método propuesto consigue llegar a un compromiso entre las técnicas de eliminación y las técnicas de auto-completado. Se consigue trabajar con la mayor información posible evitando la posibilidad de introducir errores cuando se estiman muestras.

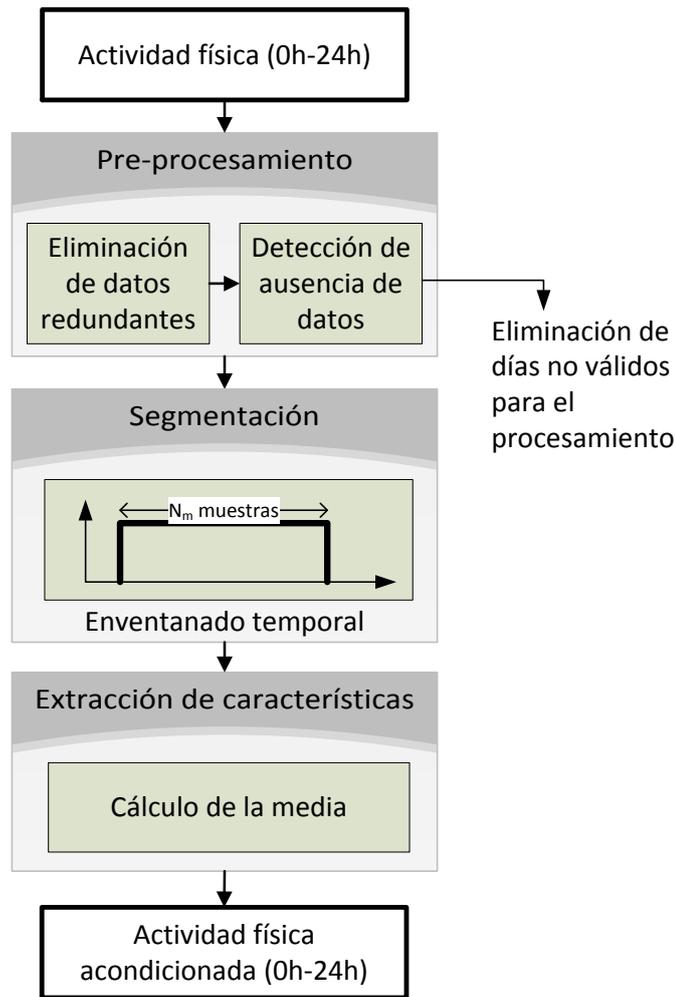


Figura 5.19. Método de acondicionamiento propuesto

```

Function calcularActividadFísicaAcondicionada(Día d, Integer  $N_{\min} = 1020$ , Integer  $N_r = 5$ , Integer  $N_m = 60$ ):
  Double[1440] ActividadFísicaAcondicionada;
  // Preprocesamiento: tratamiento de datos redundantes y datos ausentes
  preprocesarDatosRedundantes(d);
  preprocesarDatosAusentes(d,  $N_{\min}$ ,  $N_r$ );
  For Integer i from 1 to 1440:
    // Se obtiene para cada muestra i el segmento de actividad
    // de tamaño  $N_m$ 
    Double SegmentoActividad = segmentarActividadFísica(d, i,  $N_m$ );
    // Para cada segmento se obtiene la media de las medidas
    ActividadFísicaAcondicionada[i] =
      extraerCaracterísticas(SegmentoActividad);
  End for;
  Return ActividadFísicaAcondicionada;
End function;
  
```

Figura 5.20. Pseudocódigo del método de acondicionamiento propuesto

El uso de un método de segmentación basado en inventariado temporal con solapamiento total permite obtener por cada muestra de la variable un segmento que caracteriza su vecindad. El método propuesto evita la división de patrones característicos en distintos segmentos y funciona de forma automática sin necesidad de conocer la forma de la variable o los posibles patrones existentes.

Dentro de la fase de extracción de características se opta por el cálculo de la media. Esta medida de tendencia se calcula sobre cada segmento y permite obtener una medida del patrón que siguen las muestras de la variable en cada vecindad segmentada. El uso de una técnica de inventariado temporal para la segmentación (ventana cuadrada, tamaño $N_m = 60$ y solapamiento total) y la extracción de la media en cada uno de los segmentos equivale a aplicar un filtro de media sobre la variable original. Un filtro de media es un filtro paso bajo cuyo objetivo es suavizar la variable y que es de gran utilidad cuando la variable puede estar contaminada por ruido.

El método propuesto para el acondicionamiento consigue reducir el ruido presente en la variable tanto por la presencia de datos redundantes y/o ausentes como por los posibles ruidos inherentes al sensor y presentes en la medida. Esta última característica se consigue mediante la combinación de la etapa de segmentación y la etapa de extracción de características.

La Figura 5.21 muestra un ejemplo de la aplicación del método de acondicionamiento propuesto sobre una variable de actividad física. Para cada muestra n de entrada a la etapa de acondicionamiento, se tiene a la salida la media de la ventana temporal ($n - N_m/2, n + N_m/2$). Así, para cada muestra de entrada n , la salida será el nivel medio de actividad física registrado en su vecindad (N_m). Tal y como se aprecia en la Figura 5.21.b la variable de salida tendrá el mismo número de muestras que la variable original.

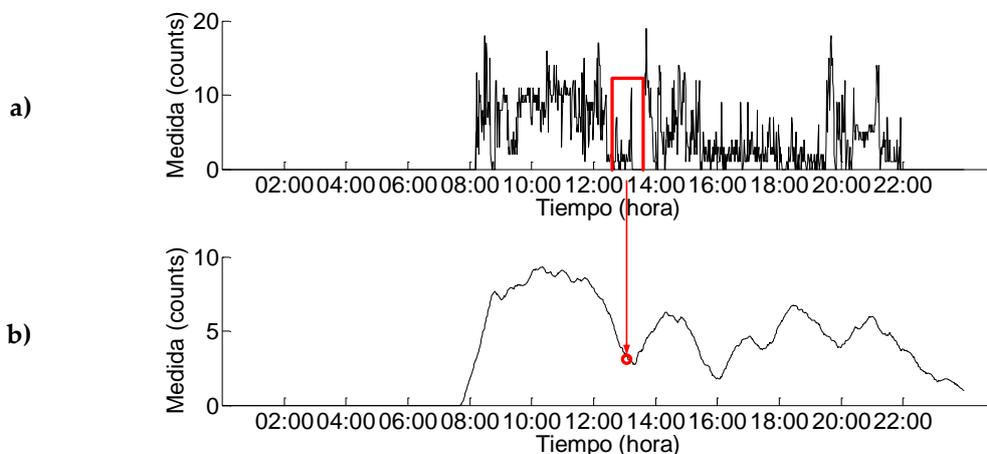


Figura 5.21. Ejemplo de variable original y variable acondicionada

La expresión matemática que define la actividad física acondicionada (salida del método de acondicionamiento) en función de la actividad física medida (salida del método de medida propuesto en el Capítulo 4), se representa en la Ecuación 5.3. Para cada muestra n de la variable de entrada, la salida es la media de la actividad física en la ventana temporal definida por el parámetro N_m . El cálculo se representa para valores de n entre 1 y 1440 que será el número de muestras medidas en un día completo de datos.

$$Actividad\ física\ acondicionada(n) = \frac{1}{N_m} \sum_{i=1+\left(n-\frac{N_m}{2}\right)}^{n+\frac{N_m}{2}} Actividad\ física(i)$$

$$n \in [1,1440]$$

Ecuación 5.3

En el Apéndice H (sección H.3) se presenta un ejemplo de aplicación del método propuesto sobre un usuario de ejemplo cuyas características se describen en la sección H.1.

5.3.3. MÉTODO DE PROCESADO (RF-5)

Tal y como ya se adelantó en la sección 5.3.1, en este trabajo se opta por un método basado en técnicas de aprendizaje automático para la obtención de un modelo de comportamiento que permita identificar la rutina de un usuario. Dentro de las técnicas de aprendizaje automático para la obtención de un modelo de comportamiento destacan principalmente los métodos estadísticos. Este tipo de métodos son versátiles, fácilmente aplicables para su trabajo con datos históricos y han demostrado su efectividad en diferentes escenarios (Kayacan et al. 2010).

Existen, como se describió en la sección 5.2.3, dos tendencias principales dentro de los métodos estadísticos: métodos paramétricos y métodos no paramétricos. Dentro del estado del arte destacan principalmente los métodos paramétricos. Estos métodos presentan una gran robustez pero el principal inconveniente en su aplicación es la asunción de que los datos siguen una determinada distribución estadística y la necesidad para ello de una exhaustiva etapa de entrenamiento. Un cambio en la distribución de las medidas, harían que fuera necesario un rediseño, reentrenamiento y reconfiguración del método. Debido a la naturaleza dinámica del comportamiento humano y a la necesidad de que el método propuesto sea adaptativo, un método paramétrico no es adecuado en el modelado.

En este trabajo se propone un método no paramétrico de procesamiento basado en el cálculo de parámetros estadísticos y en el uso de información histórica para la identificación de rutinas. El método propuesto es adaptativo retroalimentándose en su funcionamiento de los propios datos de salida. Así, se consiguen soportar los posibles cambios de rutina del usuario con el paso del tiempo sin necesidad de redefinir o reconfigurar el método.

La rutina del usuario es identificada en este trabajo mediante el cálculo del nivel de actividad física esperado de acuerdo a la información histórica existente. En este cálculo (representando en la Figura 5.22) se utiliza la actividad física acondicionada sobre los N_r últimos días para crear el modelo de comportamiento esperado. Este modelo es una variable con el mismo tamaño que las medidas de actividad originales (desde las 00:00 hasta las 24:00) y que representa la rutina esperada para el día N_r+1 . Así, para un día d , el modelo de comportamiento se calcula mediante el procesamiento de la actividad física medida en los N_r días comprendidos en $[d-1, d-N_r]$.

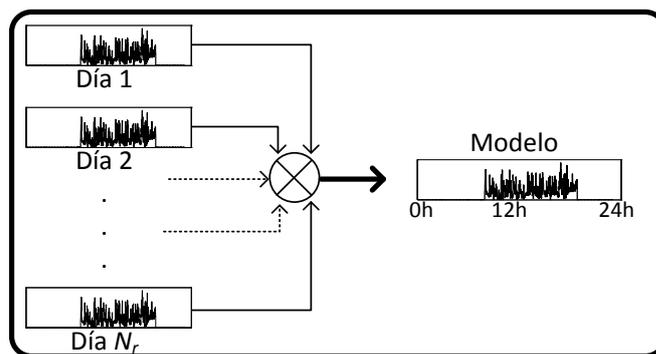


Figura 5.22. Etapa de clasificación del método de modelado propuesto

La Ecuación 5.4 muestra la expresión matemática con la que se realiza el cálculo de la rutina del usuario. Cada muestra n de la rutina para el día d se calcula como la media de la actividad física acondicionada medida en el instante n sobre los N_r últimos días. Así, el modelo de comportamiento que representa la rutina se calcula sumando y promediando muestra a muestra la salida de la etapa de extracción de características para los N_r últimos días.

$$Rutina_{Día\ d}(n) = \frac{1}{N_r} \sum_{j=(d-N_r)}^{d-1} Actividad\ física\ acondicionada_{Día\ j}(n)$$

$$n \in [1,1440]$$

Ecuación 5.4

En la etapa de acondicionamiento, a través del método de segmentación y de extracción de características propuesto, se realiza una transformación de la medida original de actividad para que cada muestra de la actividad física acondicionada represente la tendencia de la actividad medida en su vecindad. Mediante el método de procesado propuesto, se lleva a cabo la combinación de estos valores con los valores medidos durante los días anteriores. Así, la unión del método de acondicionamiento con el método de procesado hace que la rutina obtenida a la salida se construya simultáneamente a partir de información “intra-diaria” e “inter-diaria”.

La Figura 5.23 intenta representar gráficamente la característica comentada en el párrafo anterior. El método de acondicionamiento hace que cada muestra de la actividad física acondicionada represente la tendencia de la actividad física en el tramo horario en que se encuentra (información “intra-diaria”). La aplicación del método de procesado sobre la actividad física acondicionada en varios días, hace que cada muestra de rutina represente la tendencia de la actividad en los días utilizados en el procesado (información “inter-diaria”). Así, ambas características consiguen ser representadas en la salida.

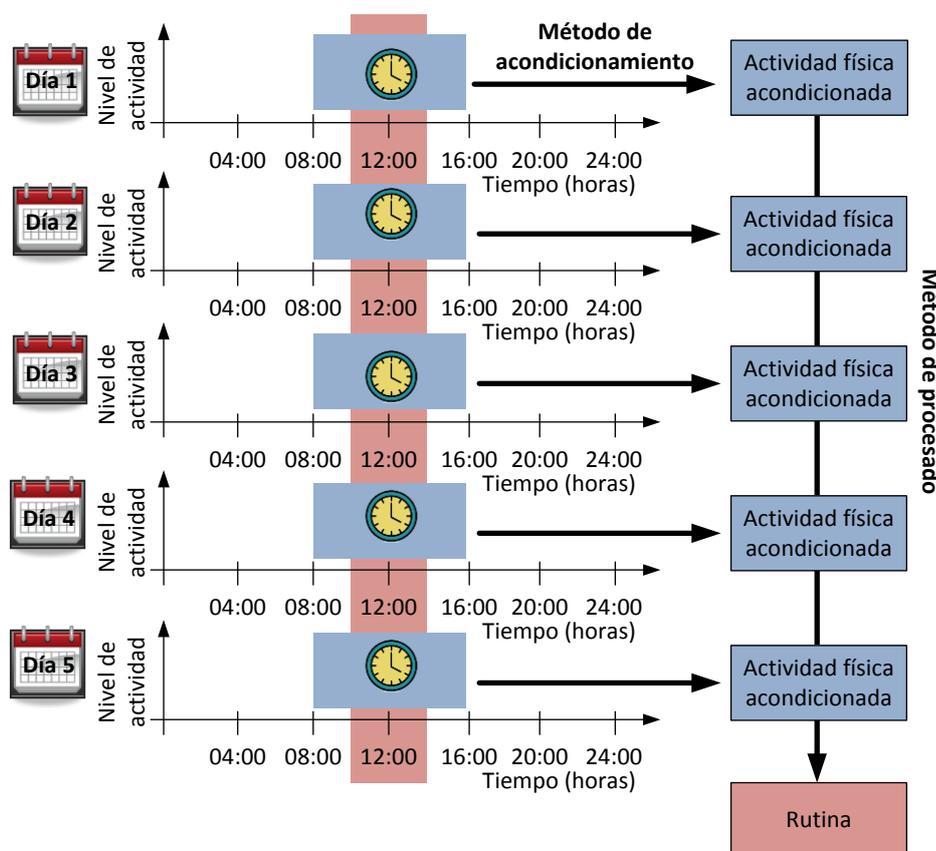


Figura 5.23. Representación gráfica del método de modelado propuesto

Al trabajar con la información histórica medida durante los N_r últimos días, el modelo que se propone es adaptativo ya que tendrán en cuenta las variaciones temporales que pueden tener lugar en el nivel de actividad del usuario. Un cambio de rutina podrá ser identificado y representado en el modelo calculado. Esta característica se consigue mediante el uso de la media, que hace que el resultado se adapte rápidamente a cambios grandes de comportamiento al tratarse de una medida sensible a la aparición de valores extremos.

La Figura 5.24 y la Figura 5.25 muestran, respectivamente, el diagrama de bloques propuesto y el pseudocódigo para el procesado de la actividad física con el objetivo de crear un modelo de comportamiento. Este modelo representa la rutina del usuario y se calcula en el día d mediante la información de actividad medida en los N_r últimos días.

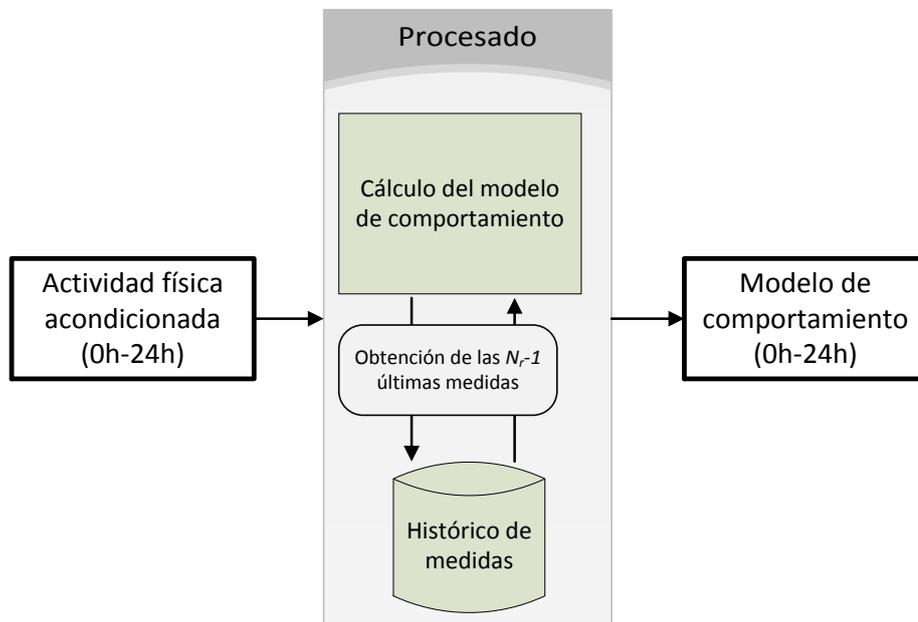


Figura 5.24. Método propuesto para el modelado del comportamiento

```

Function calcularRutina(Día d, Integer  $N_{\min} = 1020$ , Integer  $N_r = 5$ , Integer  $N_m = 60$ )
    Double[1440] Rutina;
    // Para los  $N_r$  días anteriores al día d se calcula la actividad física acondicionada
    // y se aplica el método de modelado propuesto
    For Integer i from (d- $N_r$ ) to (d-1):
        // Implementación del cálculo mostrado en la Ecuación 5.4
        Rutina += (1/ $N_r$ ) *
            (calcularActividadFísicaAcondicionada(i,  $N_{\min}$ ,  $N_r$ ,  $N_m$ ));
    End for;
    Return Rutina;
End function;
    
```

Figura 5.25. Pseudocódigo del método de modelado propuesto

El pseudocódigo utilizado para la implementación del método de modelado se representa en la Figura 5.25. La función *calcularRutina* recibe como parámetros la identificación del día para el cuál se desea calcular la rutina (d), los parámetros necesarios para la etapa de acondicionamiento (N_{min} y N_r) y el número de días empleados en el cálculo (N_r). El pseudocódigo de la función implementa la expresión matemática de la Ecuación 5.4.

Como valor de N_r se propone el uso de 5 días. Este número de días es fijado de acuerdo a las asunciones de rutina llevadas a cabo en algunos trabajos como (Heil et al. 2012) (Trost et al. 2005). Según estos trabajos, el uso de 5 días de medida para la creación de un modelo de comportamiento garantiza que este representa la rutina del usuario. Al trabajar con usuarios mayores de 60 años, este parámetro ($N_r = 5$) es adecuado ya que según (Fouquet and Franco 2010) el comportamiento rutinario en este tipo de usuarios es muy pronunciado.

Al contrario que la mayoría de los trabajos existentes en el estado del arte, el método propuesto no realiza ninguna asunción sobre la distribución estadística de los datos, ni sobre sus propiedades. El método propuesto para el cálculo de la rutina se basa solamente en los niveles medidos de actividad física, no requiere de ninguna etapa de entrenamiento previa a su funcionamiento y es capaz de adaptarse a los posibles cambios de rutina del usuario. Estas tres últimas características son los aspectos más innovadores del método propuesto.

5.3.4. VALIDACIÓN DEL MÉTODO PROPUESTO

El método propuesto para el modelado de la actividad física y obtención de la rutina del usuario es validado mediante la monitorización de 8 usuarios (4 hombres y 4 mujeres) pertenecientes a diferentes segmentos de edad. La validación sigue parte de las directrices de (Welk et al. 2012) y se ha optado por trabajar tanto con usuarios mayores de 60 años como con usuarios pertenecientes al rango de edad entre 20 y 60 años. Debido a la diferencia en los tipos de actividad desarrollados por cada conjunto de usuarios, trabajando con ambos grupos se consigue realizar una validación más amplia del método.

Cada uno de los usuarios ha sido monitorizado durante 5 días consecutivos mediante un dispositivo vestible con apariencia de reloj y acelerómetro integrado (Texas Instruments 2014). Dentro de este dispositivo se implementó el método descrito en el Capítulo 4 para la medida de la actividad física. La Tabla 5.9 muestra el resumen de las características que definen a los usuarios involucrados en el sistema y cuyos detalles pueden encontrarse en la sección E.2 del Apéndice E.

Tabla 5.9. Usuarios involucrados en la validación del método de modelado

	Media	σ
Edad	51,40 años	19,50 años
Peso	69,71 Kg	12,81 Kg
Estatura	1,66 m	0,06 m
Índice de masa corporal	24,05	4,32

En el Apéndice G se describe el procedimiento utilizado para la experimentación. En este caso, al finalizar el periodo de monitorización (5 días) se pidió a los usuarios que cubrieran un único cuestionario indicando para cada tramo horario, cual es el nivel de actividad que identifica a su rutina en el tiempo monitorizado. Tal y como se describe en la sección G.2 del Apéndice G, a partir del cuestionario de actividad cubierto se construye una variable que indica para cada tramo horario el nivel de actividad realizado. Esta variable representa la rutina teórica del usuario y es utilizada para la validación de la rutina identificada mediante el método de modelado propuesto

En la Figura 5.26 se representan los coeficientes de correlación entre el modelo calculado y la rutina teórica construida a partir de los cuestionarios de cada uno de los 8 usuarios participantes en el experimento. Esta medida de asociación ya fue utilizada en el Capítulo 4 para la validación del método de medida y es un indicador ampliamente utilizado para el estudio de la relación entre variables. Tal y como se aprecia, la correlación obtenida es superior en todos los casos a 0,85 con un valor medio de 0,92 ($\sigma = 0,03$; p-value < 0,05).

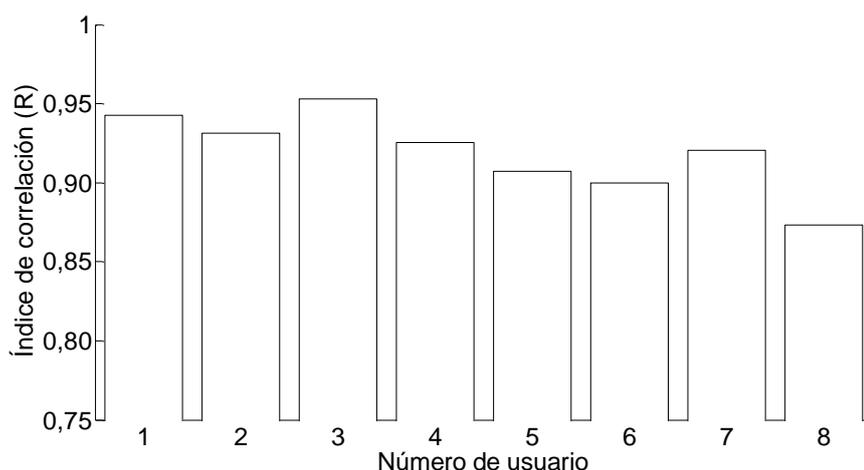


Figura 5.26. Correlación entre la rutina calculada y la especificada por los usuarios

La Figura 5.27 muestra la distribución de valores obtenidos por el modelo propuesto (*valor rutina modelo*) y los valores dados por los usuarios según los cuestionarios de actividad (*valor rutina teórica*). El diagrama de cajas mostrado muestra la distribución condicional existente entre ambas medidas. Dado un valor teórico de rutina, el diagrama de cajas correspondiente representa la distribución de los datos proporcionados por el modelo. Este gráfico permite comparar el grado de solapamiento (coincidencia) de las distribuciones condicionales. En general, cuanto mayor sea el solapamiento, menor será la intensidad de la asociación entre las dos variables y, viceversa, cuanto menor sea el solapamiento, mayor será el tamaño del efecto de la relación. En la Figura 5.27 puede apreciarse poco grado de solapamiento por lo que la intensidad de la asociación entre la medida teórica y la medida dada por el modelo es elevada.

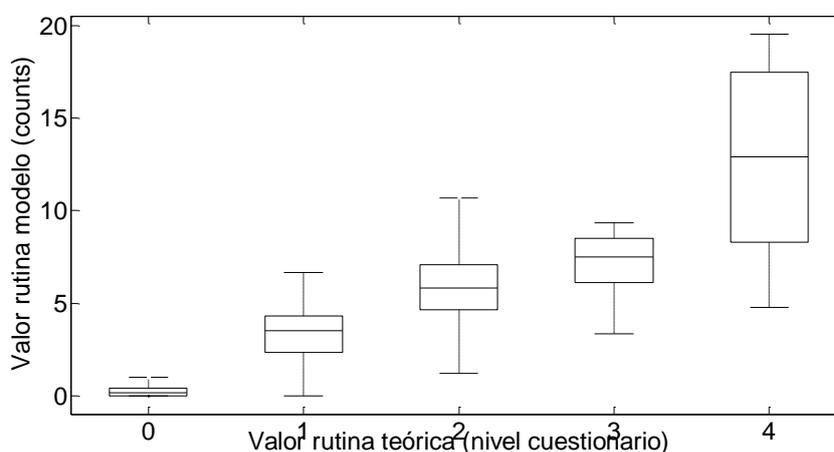


Figura 5.27. Relación entre el valor de rutina teórica y el modelo obtenido

Pese a que la medida de la rutina teórica se interpreta en este trabajo como una variable cuantitativa de valores discretos, en realidad se trata de una variable categórica de cinco categorías: actividad nula, actividad baja, actividad media, actividad alta y actividad muy alta.

En (Pérez 2005) para el análisis de la asociación entre variables cuantitativas y variables categóricas no dicotómicas (más de dos categorías), en el campo del comportamiento, se propone el uso del índice f de Cohen calculado a partir de la expresión mostrada en la Ecuación 5.5. En esta expresión Y representa la variable cuantitativa (modelo calculado), k cada una de las cinco categorías existentes en la rutina teórica y n el número de observaciones. Igual que en el caso del índice de correlación, este índice toma valores entre 0 y 1 mostrando una mayor asociación cuando más cercanos a la unidad sean estos valores.

$$f = \frac{\sigma_{\bar{Y}}}{\sigma_Y} \text{ con } \sigma_{\bar{Y}} = \sqrt{\frac{\sum_{i=1}^k n_i \cdot (\bar{Y}_i - \bar{Y})^2}{n}}$$

Ecuación 5.5

El índice f de Cohen toma un valor de 0,88 en el análisis realizado probando de nuevo la asociación existente entre el modelo calculado y la rutina teórica. De esta forma se valida de nuevo el método de modelado propuesto en este trabajo.

El método propuesto en este capítulo para el modelado del comportamiento humano permite, como se acaba de demostrar, identificar la rutina del usuario. Se demuestra la relación existente entre el modelo obtenido y la rutina teórica indicada. El método de modelado propuesto es por tanto capaz de representar correctamente la rutina del usuario a partir del procesamiento de la actividad física medida durante los últimos días.

5.3.5. RESUMEN

En el método de modelado propuesto en este capítulo intervienen tres parámetros que permiten realizar su configuración: N_{min} , N_m y N_r . La descripción de estos parámetros y sus valores propuestos se muestran en la Tabla 5.10.

Tabla 5.10. Resumen de los parámetros implicados en el método de modelado

Parámetro	Valor propuesto	Descripción
N_{min}	1020	Mínima cantidad de datos utilizada en el método de acondicionamiento para determinar si debe utilizarse una estrategia de eliminación o de auto-completado
N_m	60	Número de muestras relacionado con la resolución que se desea en el método
N_r	5	Número de días utilizado para el cálculo de la rutina

La Tabla 5.11 parte de los requisitos funcionales que debe cumplir un método de procesamiento del comportamiento y que fueron descritos en la sección 2.3.3. Se exponen las decisiones tomadas en este capítulo y la forma en la que se cumplen los requisitos establecidos.

Tabla 5.11. Requisitos funcionales del método de procesamiento

Referencia	Requisito	Descripción
☑ RF-4	Método de acondicionamiento	<p><i>Se proponen:</i></p> <ul style="list-style-type: none"> - Técnicas de pre-procesamiento para la eliminación de datos redundantes y el tratamiento de datos ausentes. - Técnica de segmentación basada en un enventanado móvil con solapamiento total. - Técnica de extracción de la media para cada segmento <p><i>La salida es una variable del mismo número de muestras que la actividad física de entrada pero con un valor que representa la tendencia de actividad medida en la vecindad de la muestra</i></p>
☑ RF-5	Método de modelado	<p><i>Se propone un método basado en métodos estadísticos que funciona de forma adaptativa y automática mediante el uso de información histórica</i></p>
☑ RF-6	Modelo esperado	<p><i>El modelo esperado representa adecuadamente la rutina del usuario y permite conocer la tendencia de actividad física que posee para cada instante temporal tal y como se ha demostrado en la validación realizada en la sección 5.3.4</i></p>

En la Figura 5.28 se muestran las etapas del método de procesamiento propuesto de acuerdo a las etapas descritas en la Figura 5.2. La actividad física es medida a través de un dispositivo vestible cuya variable es la entrada al método de procesamiento. El procesamiento se realiza a nivel de un día de actividad (medidas desde las 00:00 hasta las 24:00, variable de 1440 medidas de actividad). Tras realizar el pre-procesamiento eliminando datos redundantes y completando días con datos faltantes, se lleva a cabo una segmentación para la extracción de la media en la siguiente etapa del método. La idea es sustituir cada muestra de la variable original por la media de su vecindad (ventana de N_m muestras). La variable procesada de salida del método será una variable con el mismo número de medidas que la variable original calculada a través del histórico de medidas durante los N_r últimos días medidos.

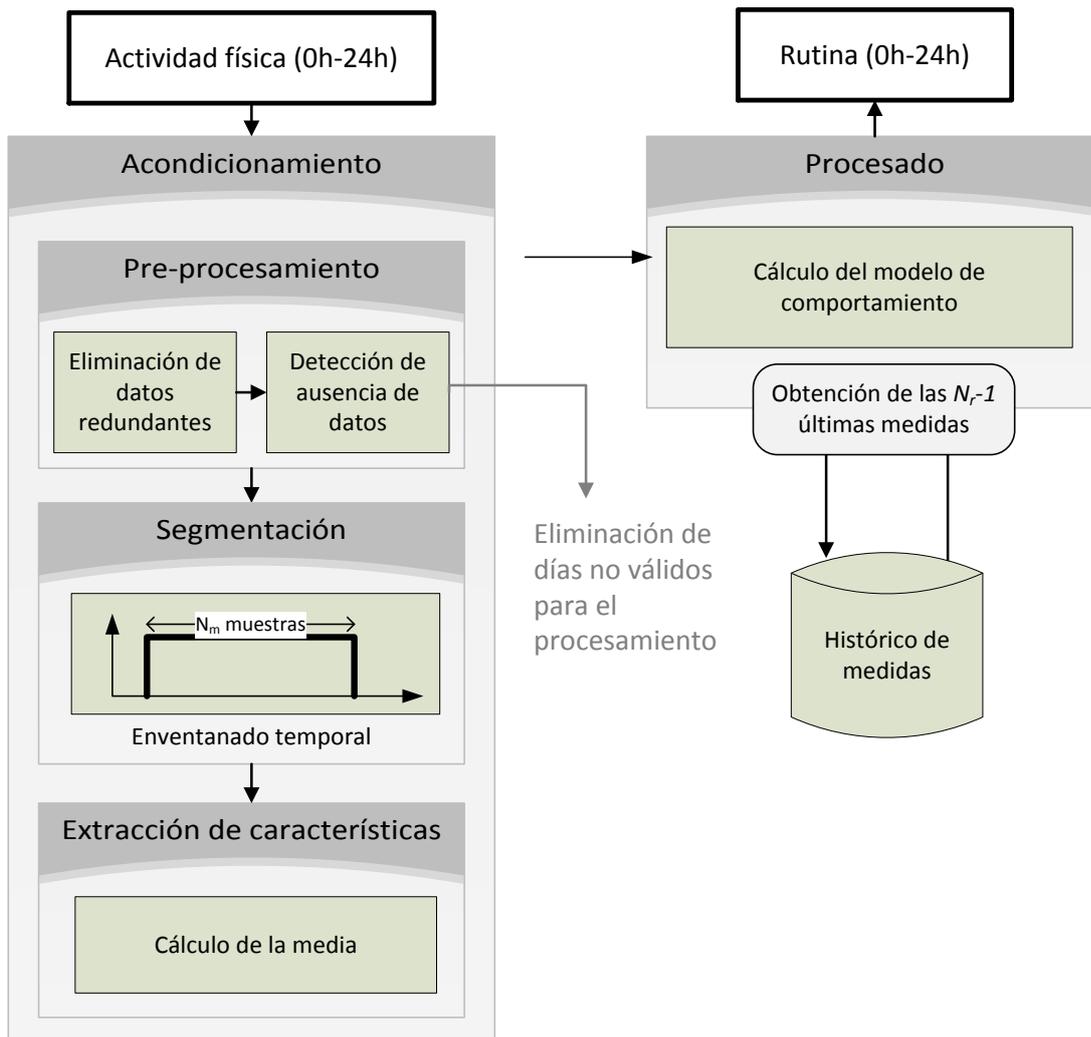


Figura 5.28. Etapas del método de modelado propuesto

5.4. CONCLUSIONES

Las principales tendencias existentes en el estado del arte para el modelado del comportamiento humano se basan en técnicas de reconocimiento de actividad. Las técnicas basadas en el reconocimiento de actividad se basan en la asunción de que el comportamiento puede dividirse en una secuencia de actividades que tiene lugar de forma consecutiva. En la realidad, las actividades pueden tener lugar de forma concurrente o intermitente. A su vez, este tipo de técnicas requieren de una etapa previa de entrenamiento y etiquetado que introduce errores y sesgo humano en el sistema.

Otra de las tendencias existentes es el uso de métodos de aprendizaje automático. Dentro de los métodos de aprendizaje automático para el modelado del comportamiento destacan los métodos estadísticos. Este tipo de métodos permiten obtener buenos resultados pero su funcionamiento se basa generalmente en asunciones acerca de la distribución de los datos de medida (métodos paramétricos). Estos métodos no permiten la detección automática y adaptativa de la rutina del usuario ya que un cambio en la distribución de los datos implicaría un rediseño y reconfiguración del método.

En este trabajo se propone un método de modelado que permite identificar la rutina del usuario de forma automática y adaptativa. Se trata de un método estadístico no paramétrico basado en datos históricos. El método propuesto se basa únicamente en la medida de la actividad física medida durante un conjunto determinado de días (N_r días) y obtiene una variable capaz de representar la rutina durante el conjunto de días estudiado. El método diseñado es dinámico y fácilmente ajustable a los cambios que se producen en el comportamiento humano ya que solamente depende de las medidas tomadas. Para cada instante se obtiene un valor de nivel de actividad esperado que depende tanto de medidas "intra-diarias" (dentro del propio día) como de medidas "inter-diarias" (medidas de los días anteriores).

El método propuesto para el modelado del comportamiento ha sido validado en un entorno real a través de la monitorización de 8 usuarios de diferentes perfiles de actividad durante 5 días. Para cada usuario se ha estudiado la asociación entre la rutina identificada y la rutina real a través del índice de correlación (valor 0,92; p-value < 0,05), análisis de la distribución de valores y mediante el índice f de Cohen (valor 0,88). El análisis realizado ha demostrado la validez del método propuesto y cómo éste a partir de datos históricos es capaz de identificar la rutina del usuario de forma automática y adaptativa.

CAPÍTULO 6.

MÉTODO DE DETECCIÓN: ANOMALÍAS DE COMPORTAMIENTO

6.1. INTRODUCCIÓN

A partir de los niveles de actividad física, en el capítulo anterior se propuso un método de modelado que permite identificar rutinas. Esa rutina, se interpreta en este capítulo como el comportamiento esperado y se propone su análisis para la detección de anomalías. A nivel general, la detección de anomalías hace referencia a la búsqueda de patrones de datos que no se adecúen a lo normal o esperado (Chandola et al. 2009) (Estevez-Tapiador et al. 2004). Cualquier anomalía puede ser traducida en información significativa y dependiendo de las circunstancias, crítica.

En el campo del comportamiento humano, las anomalías pueden ser definidas como cambios en el nivel de actividad del usuario no esperados de acuerdo a su nivel usual de actividad (Lühr et al. 2007) (Antonakaki et al. 2009). Una anomalía de comportamiento hace referencia a la detección de niveles de actividad no esperados según la rutina definida. Por ejemplo, si la rutina de una persona se caracteriza por la realización de actividades deportivas todas las tardes, se considerará una anomalía que esa persona comience a mantener un comportamiento sedentario durante el tramo horario en el que solía hacer deporte. La detección de anomalías es útil para la

detección y monitorización de problemas de salud como Alzheimer, demencia e incluso enfermedades pulmonares.

El objetivo de este capítulo es el “Diseño y validación de un método de detección de anomalías” tal y como se describe gráficamente en la Figura 6.1. Siguiendo las etapas descritas en la Figura 2.6 (acondicionamiento y evaluación), en este capítulo se parte de la magnitud seleccionada en el Capítulo 4 (actividad física) y de la rutina identificada en el Capítulo 5 (rutina). Ambas variables deben ser procesadas mediante un método de acondicionamiento (RF-7) para su posterior estudio mediante un método de evaluación (RF-8). El método de evaluación ha de ser capaz de detectar anomalías de comportamiento (RF-9) a partir de la rutina calculada y de la actividad física medida.

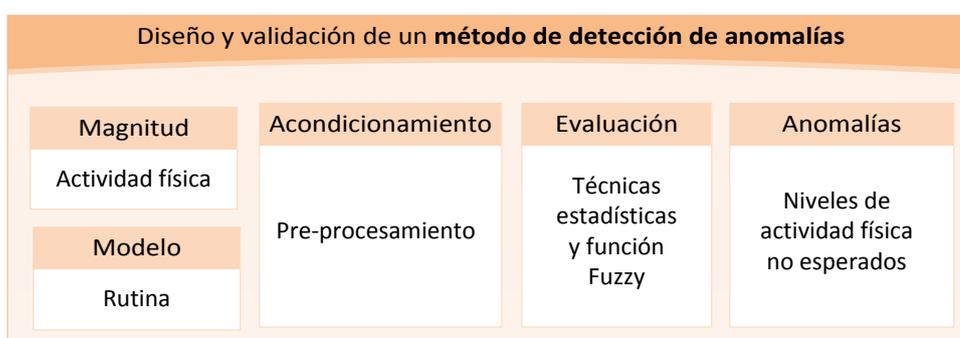


Figura 6.1. Objetivo principal del Capítulo 6

En la primera parte de este capítulo (sección 6.2) se describen los principales métodos existentes en el estado del arte para la detección de anomalías: métodos no supervisados, métodos supervisados y métodos semi-supervisados. Para cada uno de estos métodos se estudian sus ventajas e inconvenientes y su aplicabilidad a la detección de anomalías de comportamiento de acuerdo a los requisitos establecidos en este trabajo (detección automática y adaptativa). En esta sección se identifican también las principales tendencias en detección de anomalías de comportamiento. En el estado del arte existen varios trabajos centrados en ámbitos muy específicos y con objetivos muy concretos. Un ejemplo son los trabajos centrados en la detección de caídas o en la detección de trastornos específicos (insomnio, hiperactividad, demencia). Este tipo de trabajos han conseguido buenos resultados dentro de su ámbito pero ninguno de ellos permite la detección genérica de anomalías.

En la segunda parte del capítulo (sección 6.3) se propone el diseño de un método de detección de anomalías de comportamiento que pueda ser utilizado de forma general y que permita cierto grado de configuración para su adecuación a trastornos específicos. Dentro de la clasificación presentada en la sección 6.2, el método propuesto se puede clasificar como semi-supervisado consiguiendo así una buena tasa de detección con

una baja complejidad en la implementación. En este tipo de métodos se asume la existencia de un modelo que caracteriza el comportamiento esperado (rutina) y en base a él se lleva a cabo la detección de anomalías.

El método propuesto se basa en la combinación de una medida de similitud junto con técnicas de lógica difusa para la evaluación del comportamiento anómalo. Para cada medida de actividad instantánea se obtiene un valor que identifica su grado de anomalía. A la salida del método, un proceso de filtrado permitirá en función del contexto determinar comportamientos anómalos significantes. Este proceso de filtrado permite personalizar el método en función del usuario y del ámbito en el que se esté trabajando.

6.2. ESTADO DEL ARTE: DETECCIÓN DE ANOMALÍAS DE COMPORTAMIENTO

6.2.1. INTRODUCCIÓN

Formalmente, la detección de anomalías se define en (Esling and Agon 2012) para una serie temporal $T = (t_1, \dots, t_n)$ como la detección de todas las sub-secuencias T' que no cumplan el modelo esperado. En este trabajo, el modelo esperado será la rutina identificada y la actividad física medida, la serie temporal a analizar. Así, la detección de anomalías se referirá para la actividad física medida durante el periodo T , como la detección de todas las sub-secuencias T' que no se adecúen a la rutina identificada.

Un método de detección de anomalías puede solamente identificar las muestras anómalas o simultáneamente proporcionarles una puntuación de anomalía. Los métodos de análisis que proporcionan una puntuación indican para cada instancia de datos cuál es su grado de anomalía. Así, dada una serie temporal T , se podría obtener una lista T'' ordenada de anomalías en función de su puntuación.

La detección de anomalías tiene diversas aplicaciones en diferentes ámbitos y con diferentes objetivos. En (Chandola et al. 2009) se destaca la detección de anomalías y se describen sus particularidades en diferentes campos (vigilancia, medicina, procesos industriales). Dentro del estudio del comportamiento humano, la Tabla 6.1 muestra las principales tendencias existentes en el estado del arte. Existen pocos trabajos que se centren de forma genérica en la detección de anomalías de comportamiento sin centrarse en ningún ámbito o enfermedad concreta.

Tabla 6.1. Detección de anomalías en el estado del arte

Objetivo	Descripción	Referencias
Detección de caídas	Se trata del tipo de análisis más extendido en el estado del arte. Es necesaria una monitorización en tiempo real de la actividad del usuario.	(Hajj Chehade et al. 2012) (Boissy et al. 2007) (Mubashir, Shao, and Seed 2013)
Detección de inactividad	Detección de periodos anómalos de inactividad. La monitorización puede ser o no en tiempo real.	(Cuddihy et al. 2007) (Reilly et al. 2003) (Taraldsen et al. 2012)
Monitorización del sueño	Detección de cambios de comportamiento nocturnos. La monitorización se realiza a nivel diario.	(Weiss et al. 2010) (Gil et al. 2013)
Monitorización de enfermedades	Detección de anomalías propias de trastornos asociados a determinadas enfermedades como Alzheimer, demencia o COPD. La monitorización en tiempo real puede ser requerida en función de la anomalía a monitorizar.	(Lotfi et al. 2011) (Virone 2009) (Mahmoud 2013) (Waschki et al. 2012) (Dias et al. 2012) (Bernabei et al. 2010)

Existen tres tipos principales de métodos para la detección de anomalías en función de la disponibilidad de datos etiquetados (Hodge and Austin 2004): métodos no supervisados, métodos supervisados y métodos semi-supervisados. La Figura 6.2 representa gráficamente estos métodos, sus principales características y la asunción de la que parten.

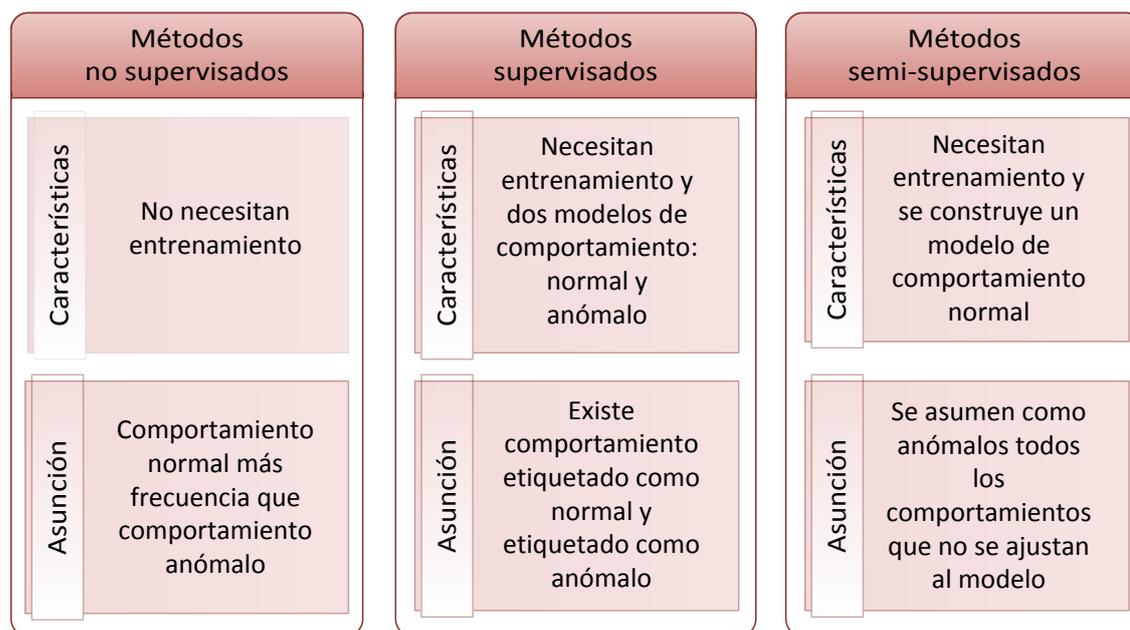


Figura 6.2. Principales métodos para la detección de anomalías

6.2.2. MÉTODOS NO SUPERVISADOS

Los métodos no supervisados no requieren datos de entrenamiento y se basan generalmente en la asunción de que el comportamiento normal es mucho más frecuente que el comportamiento anómalo. Partiendo de esta premisa, en este tipo de métodos se identifican como anomalías aquellos datos o patrones de datos poco frecuentes en las medidas.

El uso de un método no supervisado para la detección de anomalías implica la eliminación de la etapa de modelado del comportamiento descrita en el Capítulo 5 ya que las anomalías serán detectadas con los propios datos de medida sin necesidad de generar un modelo previo de comportamiento esperado (rutina). La Figura 6.3 representa el diagrama de funcionamiento de un método no supervisado para la detección de anomalías.

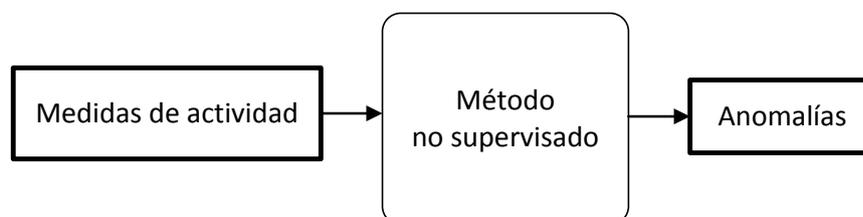


Figura 6.3. Método de detección de anomalías no supervisado

Dentro de los métodos no supervisados para la detección de anomalías destacan principalmente las técnicas de agrupamiento o *clustering*. El objetivo de estas técnicas es encontrar dentro de una serie, grupos (*clusters*) homogéneos de muestras. Se busca realizar un agrupamiento de las muestras maximizando la diferencia entre muestras de diferentes grupos y minimizando la diferencia de las muestras que se encuentran dentro del mismo.

El uso de técnicas de agrupamiento para la detección de anomalías puede abordarse según (Chandola et al. 2009) desde tres perspectivas diferentes en función de la asunción en la que se base la detección. Las siguientes tablas (de Tabla 6.2 a Tabla 6.4) presentan las principales estrategias para la detección de anomalías mediante técnicas de agrupamiento: detección basada en pertenencia, detección basada en distancia y detección basada en densidad. Se presenta y describe cada una de las estrategias junto con sus principales inconvenientes y una descripción gráfica de su funcionamiento.

La principal ventaja del uso de técnicas de agrupamiento para la detección de anomalías es la ausencia de una etapa de entrenamiento previa a la detección. Sin embargo, este tipo de técnicas no se encuentran optimizadas para la detección de anomalías sino para el agrupamiento de datos y algunos de los algoritmos son difícilmente adaptables.

Tabla 6.2. Técnicas de agrupamiento: detección basada en pertenencia

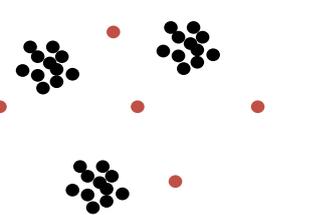
Estrategia	Asunción
Detección basada en pertenencia	Los datos normales pertenecen a alguno de los grupos existentes mientras que las anomalías no pertenecen a ninguno de ellos
Descripción	
Tras la aplicación de una técnica de agrupamiento sobre los datos de medida, se identifican como anomalías aquellas muestras que no pertenecen a ninguno de los grupos detectados	
Inconvenientes	
Los algoritmos de agrupamiento están optimizados para agrupar muestras y no para detectar las muestras que no pertenecen a ningún grupo	

Tabla 6.3. Técnicas de agrupamiento: detección basada en distancia

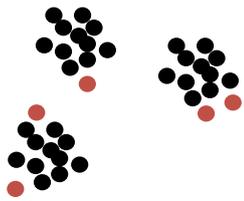
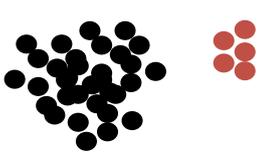
Estrategia	Asunción
Detección basada en distancia	Los datos normales se encuentran cercanos al centro de cada uno de los grupos mientras que las anomalías se encontrarán lejos de ellos
Descripción	
<p data-bbox="480 524 1086 757">Tras la aplicación de una técnica de agrupamiento sobre los datos de medida, se identifican como anomalías aquellas muestras que se encuentren alejadas del centro del grupo. Para ello, es necesario establecer un umbral de distancia</p> 	
Inconvenientes	
Si las muestras anómalas forman por ellas mismas un nuevo grupo de datos, el método no será capaz de detectarlas	

Tabla 6.4. Técnicas de agrupamiento: detección basada en densidad

Estrategia	Asunción
Detección basada en densidad	Los datos normales pertenecen a grupos de datos grandes y densos, mientras que las anomalías pertenecen a grupos pequeños y de poca densidad de datos
Descripción	
<p data-bbox="480 1314 1086 1547">Tras la aplicación de una técnica de agrupamiento sobre los datos de medida, se identifican como anomalías a aquellas muestras que pertenezcan a un grupo de datos pequeño o de baja densidad. Para ellos, es necesario establecer un umbral de distancia o densidad</p> 	
Inconvenientes	
Es necesario establecer umbrales para la detección y conocer la distribución de datos original	

6.2.3. MÉTODOS SUPERVISADOS

Los métodos supervisados asumen la existencia de datos etiquetados tanto referentes a comportamiento normal como a anómalo. Para ello, es necesaria realizar una etapa previa de entrenamiento identificando cada tipo de comportamiento y creando dos modelos de comportamiento: modelo de comportamiento normal y modelo de comportamiento anómalo. Para cada muestra de medida se evalúa su pertenencia a uno u otro modelo para su clasificación como comportamiento normal o comportamiento anómalo. La Figura 6.4 muestra el diagrama de bloques que representa el funcionamiento de este tipo de métodos. En la etapa de modelado de comportamiento (descrita en el Capítulo 5) tendría lugar la construcción del modelo de comportamiento normal y del modelo de comportamiento anómalo.

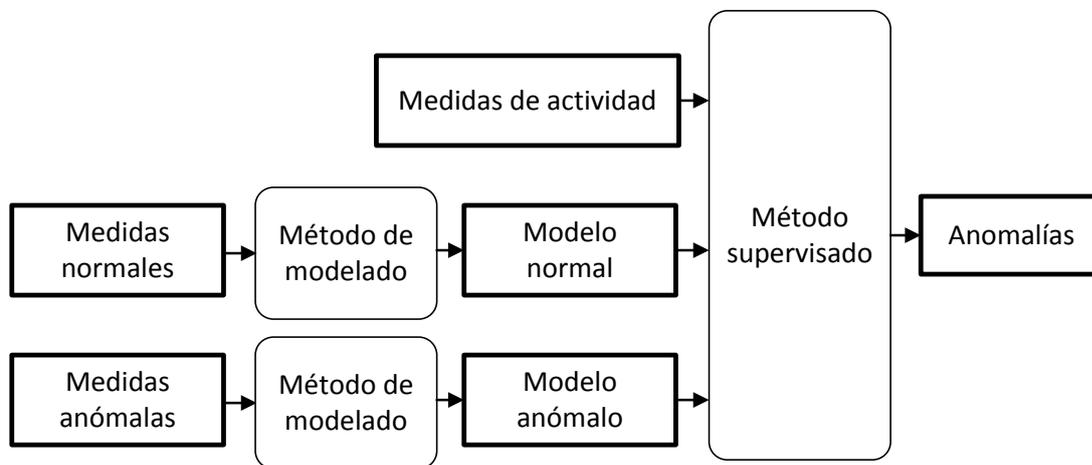


Figura 6.4. Método de detección de anomalías supervisado

La Tabla 6.5 muestra los principales métodos supervisados utilizados en el estado del arte para la clasificación y detección de anomalías. Para cada uno de ellos se presentan las ventajas e inconvenientes en su aplicación y se referencia su utilización en trabajos científicos. En (Gogoi et al. 2011) se destacan las redes neuronales y los árboles de decisión como las principales tendencias en el estado del arte.

Los métodos supervisados son eficientes cuando se trabaja con fuentes de datos estáticas cuyo rango y dominio está estrictamente definido (Hodge and Austin 2004). Un cambio en el rango o dominio de los datos de origen hace que sea necesario rediseñar el método de detección con una nueva etapa de entrenamiento para que se ajuste a las modificaciones ocasionadas.

Tabla 6.5. Principales técnicas de clasificación de comportamiento

Técnica	Descripción				
Árboles de decisión	Modelo de clasificación basado en reglas que sirven para representar y categorizar condiciones en base a la resolución de un problema				
	<table border="1"> <thead> <tr> <th>Ventajas</th> <th>Inconvenientes</th> </tr> </thead> <tbody> <tr> <td>- Fáciles de entender e interpretar - Flexibles y extensibles</td> <td>- Coste computacional alto - Diseño complejo</td> </tr> </tbody> </table>	Ventajas	Inconvenientes	- Fáciles de entender e interpretar - Flexibles y extensibles	- Coste computacional alto - Diseño complejo
	Ventajas	Inconvenientes			
	- Fáciles de entender e interpretar - Flexibles y extensibles	- Coste computacional alto - Diseño complejo			
Referencias					
(Yang 2009) (Hong et al. 2008)					
Técnica	Descripción				
Redes neuronales	Sistema de aprendizaje automático basado en la existencia de unidades de proceso muy simples trabajando en paralelo. La inteligencia del sistema reside en la conexión entre los procesos.				
	<table border="1"> <thead> <tr> <th>Ventajas</th> <th>Inconvenientes</th> </tr> </thead> <tbody> <tr> <td>- Robustez al ruido en las entradas - Diferentes tipos: multicapa, jerárquicas...</td> <td>- Etapa de aprendizaje exhaustiva y compleja - Necesidad de un conjunto cerrado de entradas y salidas</td> </tr> </tbody> </table>	Ventajas	Inconvenientes	- Robustez al ruido en las entradas - Diferentes tipos: multicapa, jerárquicas...	- Etapa de aprendizaje exhaustiva y compleja - Necesidad de un conjunto cerrado de entradas y salidas
	Ventajas	Inconvenientes			
	- Robustez al ruido en las entradas - Diferentes tipos: multicapa, jerárquicas...	- Etapa de aprendizaje exhaustiva y compleja - Necesidad de un conjunto cerrado de entradas y salidas			
Referencias					
(S. Khan et al. 2012) (Zhu and Sheng 2010) (Ermes et al. 2008) (Lee et al. 2010) (Gyórbíró et al. 2008)					
Técnica	Descripción				
Máquinas de soporte vectorial	Dado un conjunto de muestras de determina la clase a la que corresponden en función de las características extraídas				
	<table border="1"> <thead> <tr> <th>Ventajas</th> <th>Inconvenientes</th> </tr> </thead> <tbody> <tr> <td>- Etapa de entrenamiento y clasificación muy eficiente - Robusto para la generalización de problemas</td> <td>- Necesario seleccionar características interesantes para la clasificación - Difícilmente interpretables</td> </tr> </tbody> </table>	Ventajas	Inconvenientes	- Etapa de entrenamiento y clasificación muy eficiente - Robusto para la generalización de problemas	- Necesario seleccionar características interesantes para la clasificación - Difícilmente interpretables
	Ventajas	Inconvenientes			
	- Etapa de entrenamiento y clasificación muy eficiente - Robusto para la generalización de problemas	- Necesario seleccionar características interesantes para la clasificación - Difícilmente interpretables			
Referencias					
(Qian et al. 2010) (Fleury et al. 2010)					

Teniendo en cuenta que generalmente la presencia de comportamientos anómalos es menor que la existencia de comportamientos normales y que su etiquetado es complicado y poco representativo, este tipo de métodos no son generalmente utilizados en el estado del arte. Otro de los inconvenientes de los métodos supervisados es la necesidad de disponer datos etiquetados de todos los tipos de anomalías que pueden existir o estar presentes en los datos. Esto no suele ser posible debido a la naturaleza dinámica y generalmente aleatoria de las anomalías que conlleva la aparición de nuevos tipos de forma automática y difícilmente identificable (Gogoi et al. 2011) (Hodge and Austin 2004).

Así, los métodos supervisados son descartados en este trabajo para la detección de anomalías al no poder ser implementados para su funcionamiento de forma automática y adaptativa.

6.2.4. MÉTODOS SEMI-SUPERVISADOS

Los métodos semi-supervisados parten de la existencia de datos de entrenamiento etiquetados con el comportamiento normal. A partir de estos datos se construye un modelo del comportamiento esperado y se asumen como anomalías los datos que no conforman el modelo construido. Este tipo de métodos pueden trabajar de forma adaptativa mejorando y modificando el modelo a medida que llegan nuevos datos (Hodge and Austin 2004). Estos métodos se alinean con los requisitos de este trabajo y con el método de modelado diseñado en el Capítulo 5. La Figura 6.5 muestra el diagrama de bloques del funcionamiento de un método de detección semi-supervisado.

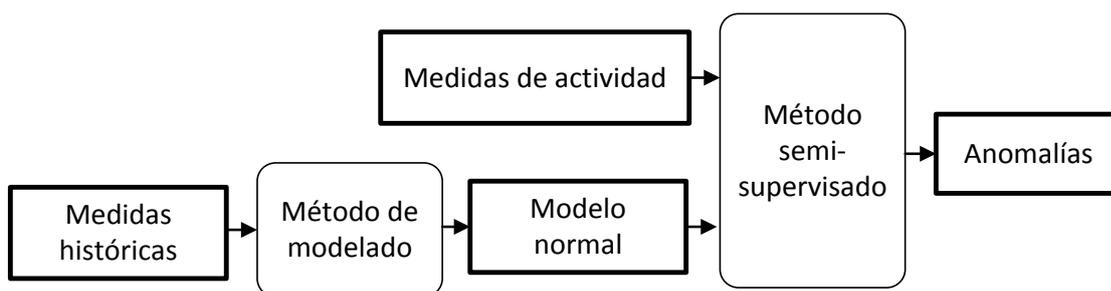


Figura 6.5. Método de detección de anomalías semi-supervisado

Dentro de los métodos semi-supervisados para la detección de anomalías destacan principalmente los métodos estadísticos cuyo uso es altamente adecuado cuando se trabaja con datos cuantitativos. Este tipo de técnicas parten de un modelo de comportamiento que se asume como normal y detectan como anomalías las muestras

que difieren estadísticamente del mismo. Existen diferentes aproximaciones en el estado del arte que pueden ser aplicadas para el análisis del comportamiento:

- *Ajuste de datos a una distribución conocida.* Esta estrategia implica el uso de métodos paramétricos para el modelado a través del ajuste de los datos a una distribución conocida. En estos casos, las anomalías pueden ser puntuadas en función de su densidad de probabilidad. Cuánta menor sea la probabilidad de aparición de un determinado valor, mayor deberá ser su puntuación de anomalía. La principal ventaja de estos métodos es el uso de técnicas estadísticas robustas y ampliamente aceptadas pero es necesario definir a priori la distribución que siguen los datos.
- *Cálculo de histograma.* A partir del cálculo del histograma que siguen los datos históricos de medida, la detección de anomalías tiene lugar en base a la relación entre la muestra a evaluar y el histograma calculado. Cada medida es evaluada en función de su frecuencia de aparición. Esta técnica es similar a la anterior pero sin necesidad de ajustar los datos a una distribución de datos conocida.
- *Diagrama de cajas.* Cuando no se realiza el ajuste de los datos a una distribución conocida, la detección de anomalías puede realizarse a través de los propios datos de medida mediante herramientas de estadística descriptiva como la construcción de un diagrama de cajas. La construcción de este diagrama se basa en el cálculo de percentiles y en la identificación como anomalías a las muestras que se encuentren fuera de un determinado rango.
- *Medidas de similitud.* Una vez establecido un modelo, la detección de anomalías puede tener lugar mediante el uso de medidas de similitud que valoren la relación entre las muestras medidas y un modelo establecido. En este tipo de técnicas, la medida de similitud utilizada es uno de los aspectos más importantes del diseño (Popoola 2012) destacando en el estado del arte el cálculo de residuos y el uso de la distancia euclídea.

6.2.5. RESUMEN

A la hora de diseñar un método para la detección de anomalías en el comportamiento humano, hay tres puntos clave que se deben tener en cuenta según (Gogoi et al. 2011):

- *Definición de datos normales.* Dependiendo de la aplicación y del ámbito de trabajo, definir el rango de datos normales no es siempre sencillo. En el caso de la medida del comportamiento humano a través de la actividad física, el rango no puede ser definido de forma universal. El conjunto de datos considerado como normales es altamente dependiente del usuario.

- *Definición de anomalías.* Es necesario disponer de una definición adecuada de anomalía en el contexto de trabajo. En el caso del análisis del comportamiento, se entenderá por anomalía a aquellos niveles de actividad física no esperados de acuerdo al comportamiento histórico del usuario.
- *Disponibilidad de datos de entrenamiento.* Dependiendo del método utilizado para la detección, es necesario disponer de datos etiquetados de comportamientos normales y anómalos tanto para el entrenamiento como para la validación del método propuesto.

La Tabla 6.6 muestra el resumen de ventajas e inconvenientes de los tres tipos de métodos identificados y estudiados en esta sección: métodos no supervisados, métodos supervisados y métodos semi-supervisados.

Tabla 6.6. Comparación de métodos para la detección de anomalías

Método	Ventajas	Inconvenientes
Métodos no supervisados	<ul style="list-style-type: none"> - No requiere entrenamiento - Fácilmente adaptables - Implementación sencilla y rápida - Detectan anomalías de forma automática y adaptativa 	<ul style="list-style-type: none"> - Asumen que los datos normales son más frecuentes que los anómalos (alta tasa de falsas alarmas) - Altamente dependientes de la eficiencia del algoritmo seleccionado - No optimizadas para la detección de anomalías - Computacionalmente complejos
Métodos supervisados	<ul style="list-style-type: none"> - Algoritmos potentes - Detección rápida - Tasa alta de detección de anomalías 	<ul style="list-style-type: none"> - Etapa de entrenamiento necesaria: datos normales y datos anómalos - Es necesario trabajar con todos los posibles datos de medida
Métodos semi-supervisados	<ul style="list-style-type: none"> - No requiere entrenamiento ni etiquetado de datos - Pueden funcionar de forma adaptativa 	<ul style="list-style-type: none"> - Se asume un conjunto de datos de entrada como normales - La tasa de falsos positivos puede ser muy alta

6.3. DISEÑO DEL MÉTODO DE DETECCIÓN DE ANOMALÍAS

6.3.1. ANOMALÍAS DETECTADAS (RF-9)

El método utilizado para la detección de comportamientos anómalos debe detectar dentro de un día (día evaluable) periodos de tiempo en los cuales el comportamiento del usuario es significativamente diferente respecto al modelo de comportamiento (rutina).

El método de detección de anomalías utilizado en este trabajo puede ser clasificado como semi-supervisado. Este tipo de métodos asumen la existencia de un modelo de comportamiento normal (rutina) y permiten identificar de forma automática y adaptativa muestras que difieren de lo esperado. Al contrario que los métodos supervisados, no requieren ni de entrenamiento ni de etiquetado de datos. Su complejidad computacional es menor que la de los métodos no supervisados y son mucho más eficientes.

El método propuesto en este trabajo para la detección de anomalías, al igual que el método utilizado para el modelado, se basa en el uso de técnicas estadísticas sin realizar asunciones sobre la naturaleza de los datos. Este método supone un avance respecto a los métodos semi-supervisados existentes en el estado del arte. La principal característica del método es la combinación de técnicas estadísticas con técnicas de minería de datos para obtener a la salida una variable continua en la que se gradúan los tramos temporales de comportamiento anómalo. Así, la salida del método propuesto presenta valores comprendidos entre 1 y -1 indicando el grado (alto o bajo) y sentido (exceso o defecto de actividad) de la anomalía.

La Figura 6.6 muestra un ejemplo de variable de comportamiento anómalo que representa las anomalías detectadas por el método. Para cada instante temporal se muestra un valor comprendido entre -1 y 1 indicando el grado y sentido de las anomalías. Un valor 0 se corresponderá con el comportamiento identificado como normal, un valor en $(0, 1]$ con un incremento anómalo de actividad y un valor en $[-1, 0)$ con un decremento anómalo de actividad. En este ejemplo, se detectan tres tramos principales de comportamiento anómalo: incremento anómalo de actividad entre las 11:00 y las 13:00, decremento anómalo de actividad en torno a las 17:00 y decremento anómalo de actividad en torno a las 20:00.

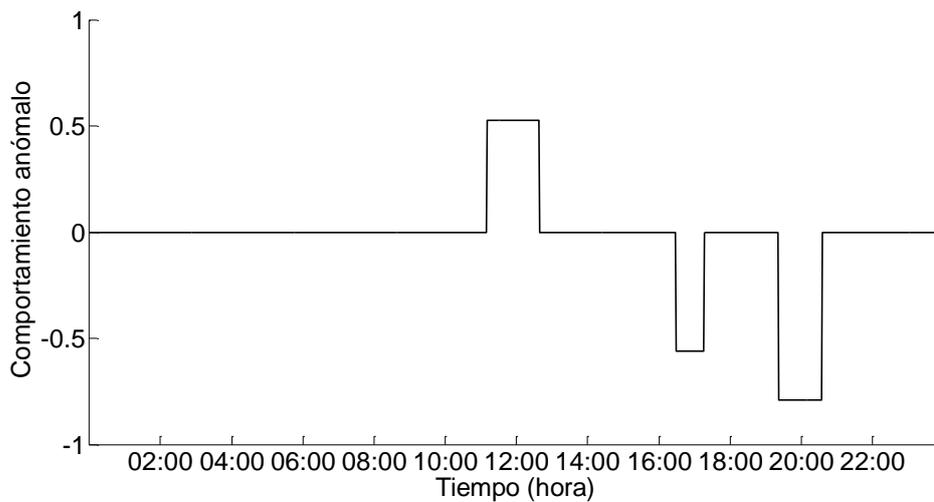


Figura 6.6. Ejemplo de anomalías detectadas

6.3.2. MÉTODO DE ACONDICIONAMIENTO (RF-7)

Las técnicas de pre-procesamiento descritas en la sección 5.3.2.2 son aplicadas sobre la medida de actividad del día evaluable (día en el que se analiza la actividad) para la eliminación de datos redundantes y datos faltantes. Este acondicionamiento de la medida permitirá mejorar su comparación con el modelo de actividad esperado.

6.3.3. MÉTODO DE EVALUACIÓN (RF-8)

6.3.3.1. INTRODUCCIÓN

En la sección F.1 del Apéndice F se presentan una serie de pruebas de concepto llevadas a cabo mediante la aplicación de los principales métodos semi-supervisados existentes en el estado del arte para la detección de anomalías de comportamiento: ajuste de datos a una distribución conocida, cálculo de histogramas, diagrama de cajas y medidas de similitud. De todos los métodos semi-supervisados de detección de anomalías presentados en la sección 6.3.2 destaca principalmente el cálculo de residuos como medida de similitud.

La técnica de ajuste de los datos a una distribución normal no es válida para la medida del comportamiento ya que los datos no se ajustan a una distribución normal. Se ha demostrado como ni los datos adquiridos ni ninguna de sus más comunes transformaciones pueden ser modeladas por este tipo de distribución. El uso de histogramas para la detección de anomalías tampoco es adecuado ya que los niveles altos de actividad tienden a presentar una baja tasa de aparición y por tanto, siempre se

van a ver puntuados con un valor alto de anomalía. Debido a la alta desviación de los datos de medida, los diagramas de cajas tampoco presentan tampoco buenos resultados.

El cálculo de los residuos entre la rutina identificada y la medida de actividad es equivalente al cálculo del valor absoluto de los residuos del modelo, tal y como se describe en la sección F.1 del Apéndice F. La ventaja del cálculo de los residuos, en lugar del uso de la distancia euclídea, es que no se trata de una medida en valor absoluto y permite conocer el sentido de la diferencia entre la rutina y la medida de actividad. Sin embargo, el principio inconveniente del cálculo de los residuos es que se trata de una medida altamente sensible a las unidades utilizadas. Las diferencias obtenidas cuando se trabaja con valores altos son generalmente mayores que las obtenidas cuando se trabaja con valores bajos y no son comparables entre sí.

El método de evaluación de anomalías propuesto en este trabajo resuelve los principales inconvenientes en el uso del cálculo de residuos para la detección de anomalías. Para resolver esta limitación, se utilizan técnicas de lógica difusa. Las técnicas de lógica difusa, también conocidas como técnicas Fuzzy, se basan en la creación de un conjunto difuso con valores comprendidos entre 0 y 1 y una función de pertenencia que permita el mapeo entre el conjunto de observaciones y el conjunto difuso definido. La función de pertenencia propuesta en este trabajo valorará los residuos del modelo obteniendo una puntuación de anomalía dependiente del instante temporal en el que se realice el cálculo.

En este trabajo se propone un método de detección de anomalías inspirado en el uso del cálculo de residuos que funciona de forma adaptativa y basado en las características propias de las variables medidas. No se realiza ninguna asunción sobre la naturaleza de la distribución que siguen los datos y genera a la salida una variable continua que indica una puntuación de anomalía para cada uno de los instantes temporales de medida. El método de detección de anomalías propuesto detecta periodos de tiempo en un día evaluable en que el comportamiento del usuario es significativamente diferente al calculado según la rutina identificada.

El método propuesto basado en el cálculo de los residuos (cálculo de la diferencia) y en el uso de técnicas de lógica difusa se divide, tal y como se representa en la Figura 6.7, en tres fases diferenciadas:

- *Cálculo de diferencia de actividad.* Cálculo de la diferencia de actividad entre la rutina y la actividad medida en el día evaluable. Esta operación es equivalente al cálculo de los residuos del modelo (rutina) para el día evaluable. El cálculo se realiza muestra a muestra entre el modelo calculado (rutina) y la medida de la actividad en el día evaluable.
- *Extracción de características estadísticas.* Cálculo de las características estadísticas que se utilizarán para la construcción de una función Fuzzy que permita la valoración de anomalías. Estas características serán extraídas tanto del modelo de actividad calculado como de la medida de actividad del día evaluable.
- *Detección de comportamiento anómalo.* A partir de los parámetros estadísticos extraídos en la fase anterior y de la diferencia de actividad calculada, la salida de esta fase será una variable continua que permita detectar muestras de actividad anómalas. A partir de estas muestras se identificarán tramos de comportamiento anómalo.

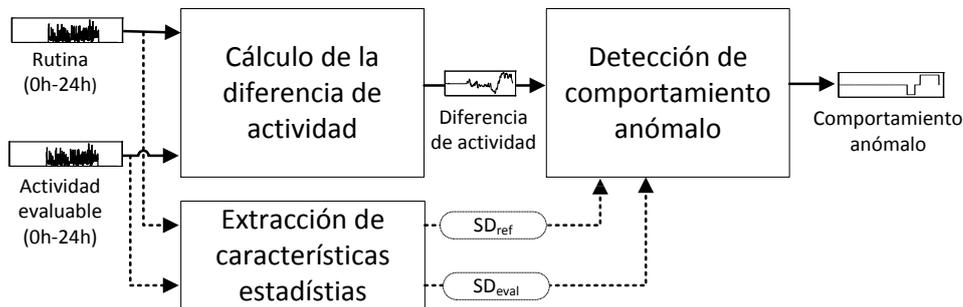


Figura 6.7. Fases del método de detección propuesto

6.3.3.2. FASE I: CÁLCULO DE DIFERENCIA DE ACTIVIDAD

El objetivo de esta fase es la obtención de una variable (*diferencia de actividad*) que represente la diferencia existente entre la actividad medida en el día evaluable (*actividad evaluable*) y el modelo de comportamiento (*rutina*). Siguiendo el análisis realizado en el Apéndice F, se utilizarán los residuos del modelo como medida de similitud y recaerá en el resto de fases del método la identificación de anomalías a partir de su cálculo.

En esta fase del método se calcula la diferencia de actividad muestra a muestra entre el modelo calculado (*rutina*) y la medida de la actividad en el día evaluable (*actividad evaluable*). La Ecuación 6.1 describe este cálculo matemático.

$$\text{Diferencia de actividad} = (\text{Actividad evaluable} - \text{Rutina})$$

Ecuación 6.1

Para cada muestra, la diferencia positiva o negativa indica la variación de actividad del día evaluable respecto a lo esperado de acuerdo con la rutina. Este modelo, como se describió en el Capítulo 5, se calcula usando la actividad medida durante los N_r días anteriores. Para eliminar el posible ruido que contamina la medida de actividad del día evaluable y suavizar el comportamiento impulsivo, se utiliza un filtrado de mediana. La implementación de este filtro asume que la variable toma valor 0 más allá de los extremos y su tamaño es de N_m muestras. El parámetro N_m puede ser interpretado como la resolución del método y está relacionado con la resolución deseada en la detección de anomalías de comportamiento.

La Ecuación 6.2 muestra la expresión matemática utilizada para la obtención de la salida de esta fase. El pseudocódigo que representa la implementación de esta fase y el diagrama de bloques que la describe pueden verse respectivamente en las Figura 6.8 y Figura 6.9.

$$\begin{aligned} \text{Diferencia de actividad filtrada} \\ = \text{filtroMediana}(\text{Actividad evaluable} - \text{Rutina}, N_m) \end{aligned}$$

Ecuación 6.2

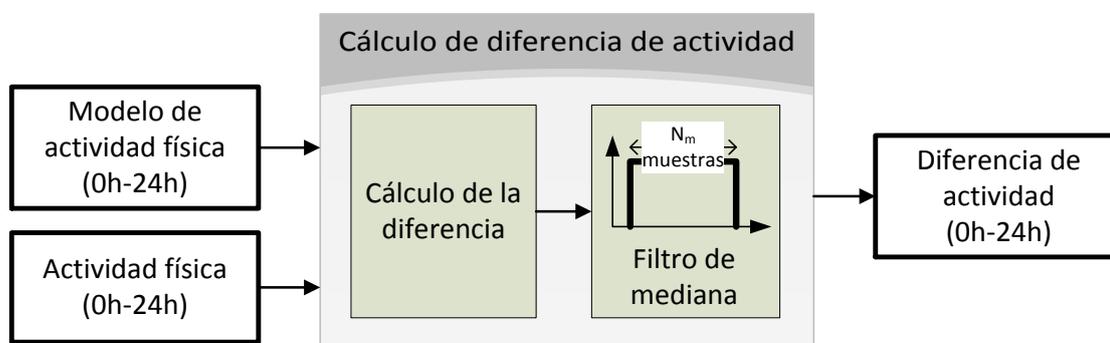


Figura 6.8. Diagrama de bloques de la fase I: cálculo de diferencia de actividad

La función *calcularDiferenciaActividadFiltrada* recibe como entrada la medida de actividad (*Actividad*) y la rutina esperada (*Rutina*) e implementa la expresión matemática mostrada en la Ecuación 6.2. El cálculo del filtro de mediana se representa en la función *filtroMediana* que recibe el parámetro N_m relacionado con la resolución deseada en la detección y que determina el tamaño de ventana del filtro utilizado.

```

Function calcularDiferenciaActividadFiltrada(Double[1440] Rutina, Double[1440] Actividad,
Integer  $N_m = 60$ ):
    // Se calcula la diferencia entre la actividad medida y la rutina
    Double DiferenciaActividad = Actividad – Rutina;
    // Se aplica filtro de mediana de tamaño  $N_m$ 
    DiferenciaActividadFiltrada = filtroMediana(DiferenciaActividad,  $N_m$ );
    Return DiferenciaActividadFiltrada;
End

Function filtroMediana(Double[] Input, Integer  $N_m = 60$ )
    Double[] Output;
    For Integer i from 1 to length(input):
        If  $N_m$  is impar:
            Output[i] = Mediana(Input[from  $i-(N_m-1)/2$  to  $i+(N_m-1)/2$ ])
        End if;
        If  $N_m$  is par:
            Output[i] = Mediana(Input[from  $i-(N_m/2)$  to  $i+(N_m/2)-1$ ])
        End if;
    End for;
    Return Output;
End function;

```

Figura 6.9. Pseudocódigo de la fase I: cálculo de diferencia de actividad

6.3.3.3. FASE II: EXTRACCIÓN DE CARACTERÍSTICAS ESTADÍSTICAS

En esta fase se realiza la extracción de características estadísticas tanto del modelo de actividad (rutina) como de la actividad a evaluar. Estas características serán utilizadas en la fase III del método para la identificación de anomalías a partir de la diferencia calculada en la fase I.

En la fase de extracción de características estadísticas, la desviación estándar (SD) es calculada tanto para el modelo de comportamiento como para la actividad medida durante el día evaluable. Para ello se realiza un inventanado temporal de tamaño N_m . Esta fase del método de detección de anomalías es similar a la etapa de extracción de características del método de modelado pero realizando el cálculo de la desviación en lugar de la media.

Para cada muestra i de la medida de actividad del día evaluable y del modelo de comportamiento, la desviación estándar es calculada respectivamente utilizando la Ecuación 6.3 y la Ecuación 6.4. Una muestra i con una desviación estándar baja indica una zona de nivel de actividad estable mientras que una desviación estándar alta indicará una zona de actividad inestable. La desviación estándar sirve para conocer la estabilidad del nivel de actividad en la vecindad de cada muestra y para realizar su caracterización no solo de acuerdo a los valores que toma sino a su dispersión.

$$SD_{rutina}(i) = DesviacionEstándar(Rutina(j; j \in [1 + i - \frac{N_m}{2}, i + \frac{N_m}{2}]))$$

Ecuación 6.3

$$SD_{evaluabile}(i) = DesviacionEstándar(Actividad\ evaluabile(j; j \in [1 + i - \frac{N_m}{2}, i + \frac{N_m}{2}]))$$

Ecuación 6.4

La Figura 6.10 muestra el pseudocódigo utilizado para la implementación de esta fase del método propuesto cuyo diagrama de bloques se representa en la Figura 6.11. La función *extraerCaracterísticasEstadísticas* recibe la variable sobre la que se va a realizar el cálculo (*Actividad*) y devuelve la desviación estándar (*SD*) calculada mediante una ventana móvil de tamaño N_m (parámetro de entrada de la función). Esta función será invocada tanto para la rutina (*Rutina*, ver Ecuación 6.3) como la actividad medida (*Actividad evaluabile*, ver Ecuación 6.4) y su salida será la salida de la fase II del método de detección propuesto.

```

Function extraerCaracterísticasEstadísticas(Double[1440] Input, Integer  $N_m = 60$ ):
    Double[1440] SD;
    // Se calcula la desviación estándar mediante una ventana móvil de tamaño  $N_m$ 
    // sobre la variable Input
    For Integer i from 1 to length(Input):
        SD[i] = calcularDesviaciónEstándar(Input[from 1 + i - ( $N_m/2$ ) to i + ( $N_m/2$ )]);
    End for;
    Return SD;
End function;

```

Figura 6.10. Pseudocódigo de la fase II: extracción de características estadísticas

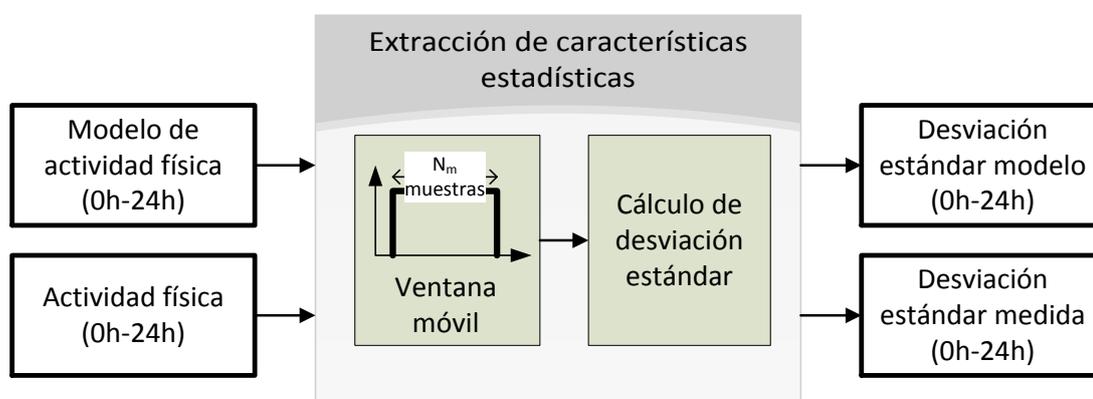


Figura 6.11. Diagrama de la fase II: extracción de características estadísticas

6.3.3.4. FASE III: DETECCIÓN DE COMPORTAMIENTO ANÓMALO

La fase de detección de comportamiento anómalo es la fase más importante del método de detección de anomalías y se basa en la comparación entre la diferencia de actividad filtrada (salida de la fase I del método) y las desviaciones estándar calculadas (salida de la fase II del método). Así, esta fase recibe como entradas las variables calculadas en las dos anteriores fases: diferencia de actividad filtrada, desviación estándar del modelo de comportamiento y desviación estándar de la medida de actividad del día evaluable.

Las secciones F.2.2 y F.2.3 del Apéndice F muestran la forma en la que las desviaciones estándar calculadas pueden ser utilizadas para la evaluación de anomalías y los resultados que cada una de estas medidas ofrecen. El uso de las desviaciones estándar calculada en la fase anterior permite establecer umbrales dinámicos y adaptativos permitiendo la detección de anomalías de forma contextual dependiendo del instante temporal de medida. El uso de la desviación estándar del modelo como medida de la dispersión esperada en la diferencia de actividad puso de manifiesto una serie de limitaciones que podían ser resueltas parcialmente utilizando la dispersión en el día evaluable.

Un compromiso entre ambos umbrales parece una de las mejores alternativas para la detección (descripción gráfica en la Figura 6.12):

- Si la diferencia de actividad es superior en valor absoluto a la desviación estándar del modelo de actividad, la actividad para ese instante temporal es anómala respecto al modelo ya que esta fuera de la dispersión de actividad esperada por el modelo en esa muestra temporal.
 - o Si la diferencia de actividad es además superior en valor absoluto a la desviación estándar sobre el día evaluable, la actividad es anómala respecto al modelo y respecto al día evaluable. La medida de actividad se encuentra fuera de la dispersión esperada tanto por el modelo como por el día evaluable.
 - o Si la diferencia de actividad no es superior en valor absoluto a la desviación estándar sobre el día evaluable, la actividad solo es anómala respecto al modelo. La medida de actividad se adecúa a la dispersión de las medidas en el día evaluable.
- Si la diferencia de actividad es inferior en valor absoluto a la desviación estándar del modelo de actividad, la actividad para ese instante temporal es normal respecto al modelo. La muestra de actividad se ajusta a la dispersión esperada por el modelo para ese instante temporal.

- Si la diferencia de actividad es superior en valor absoluto a la desviación estándar sobre el día evaluable, la actividad es anómala respecto al día evaluable. La medida de actividad se encuentra fuera de la dispersión esperada por el día evaluable.
- Si la diferencia de actividad no es superior en valor absoluto a la desviación estándar sobre el día evaluable, la actividad es normal respecto al modelo y respecto al día evaluable. La medida de actividad se adecúa a la dispersión esperada.

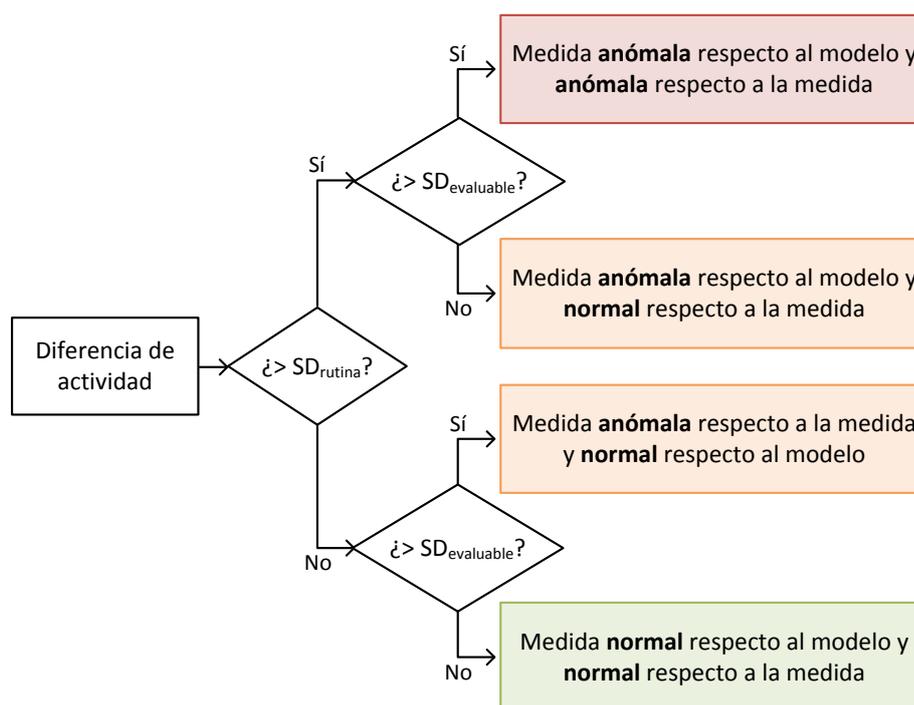


Figura 6.12. Compromiso de umbrales para detección de anomalías

La idea que se describe en los párrafos anteriores y que se representa gráficamente en la Figura 6.12 puede ser descrita matemáticamente mediante el uso de una técnica de lógica difusa. Este tipo de técnicas se basan en el establecimiento de una función de pertenencia. Esta función de pertenencia actúa sobre el conjunto X de observaciones proporcionando a la salida un valor del conjunto difuso A . A esta función de pertenencia se la suele denominar $\mu_A(x)$ y se define formalmente mediante la Ecuación 6.5.

$$\mu_A(x): X \rightarrow [0,1]$$

Ecuación 6.5

En la sección F.2.4 del Apéndice F se muestran diferentes tipos de funciones de pertenencia. Cada una de ellas presenta formas diferentes y los resultados de su aplicación sobre el conjunto de entrada serán más o menos adecuados en función de su objetivo, su ámbito de aplicación y el dominio del conjunto de entrada. Siguiendo esta idea, en este trabajo se propone la creación de una función de valoración Fuzzy de anomalías que dada la diferencia de actividad proporcione una puntuación de anomalía dependiendo de la relación de la diferencia con las desviaciones calculadas.

Las desviaciones estándar calculadas en la fase II del método son utilizadas para la construcción de una función Fuzzy basada en una función de pertenencia trapezoidal (ver Figura F.12 en el Apéndice F). La función Fuzzy planteada mapea los valores de diferencia de actividad en el conjunto difuso $[-1, 1]$ con la siguiente interpretación:

- Un valor en el conjunto $[-1, 0)$ se corresponde con la detección de anomalías por un defecto de actividad física. Se detectan muestras de actividad más bajas de lo esperado.
- Un valor en el conjunto $(0, 1]$ se corresponderá con la detección de anomalías por un incremento de la actividad física. Se detectan muestras de actividad más altas de lo esperado.
- Un valor 0 se corresponde con valores normales. Se detectan muestras de actividad dentro del rango esperado.

La Figura 6.13 muestra el aspecto de la función de pertenencia utilizada en este trabajo cuya expresión matemática se muestra en la Ecuación 6.6. Esta función, basada en una función de pertenencia trapezoidal, es definida a través de 4 parámetros característicos (a, b, c, d). Para cada muestra de diferencia de actividad x :

- Si $x < a$, el valor de anomalía será -1. La actividad realizada por el usuario es anormalmente baja.
- Si $a < x < b$, el valor de anomalía se encontrará en el rango $(-1, 0)$. El valor de anomalía será proporcional a la relación que la diferencia de actividad posee entre los valores a y b fijados (relación lineal).
- Si $b < x < c$, el valor de anomalía es 0. La actividad realizada por el usuario se encuentra dentro del rango de normalidad. Los parámetros a y b establecerán, por tanto, los rangos de diferencia de actividad física que se considerarán normales.
- Si $c < x < d$, el valor de anomalía se encontrará en el rango $(0, 1)$. El valor de anomalía será proporcional a la relación que la diferencia de actividad posee entre los valores c y d fijados (relación lineal).

- Si $x > d$, el valor de anomalía es 1. La actividad realizada por el usuario es anormalmente alta.

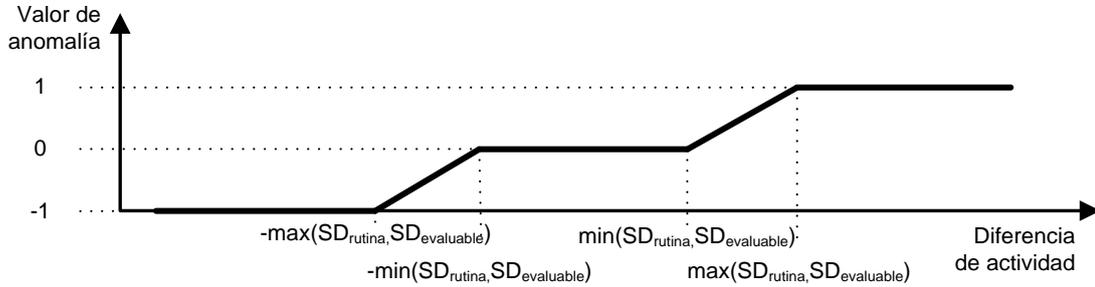


Figura 6.13. Función de valoración Fuzzy para evaluación de anomalías

$$Valor\ de\ anomalía(i) = \begin{cases} -1 & x \leq a \quad \text{Nivel de actividad muy bajo} \\ -1 + \frac{x-a}{b-a} & a < x < b \quad \text{Nivel de actividad bajo} \\ 0 & b \leq x \leq c \quad \text{Nivel de actividad normal} \\ 1 - \frac{d-x}{b-a} & c < x < d \quad \text{Nivel de actividad alto} \\ 1 & x \geq d \quad \text{Nivel de actividad muy alto} \end{cases}$$

$$x = \text{Diferencia de actividad filtrada}(i)$$

$$a = -\max(SD_{rutina}(i), SD_{evaluabile}(i))$$

$$b = -\min(SD_{rutina}(i), SD_{evaluabile}(i))$$

$$c = \min(SD_{rutina}(i), SD_{evaluabile}(i))$$

$$d = \max(SD_{rutina}(i), SD_{evaluabile}(i))$$

Ecuación 6.6

El uso de una función difusa permite graduar las anomalías en lugar de dar una valoración binaria. Así, se cumplirá el requisito de salida de poseer una variable continua que identifique el grado de anomalía detectada. El uso de ambas desviaciones es una de las características principales del método propuesto. En otros trabajos como (Chen et al. 2013; Virone 2009) las anomalías son detectadas teniendo únicamente en cuenta las características estadísticas del modelo de comportamiento.

La salida de la función Fuzzy utilizada será un valor de anomalía para cada una de las muestras i de entrada. Así, a la salida se tendrá una variable continua con valores comprendidos entre -1 (anomalía por defecto de actividad) y 1 (anomalía por exceso de actividad). Los valores comprendidos entre -1 y 0 simbolizarán una anomalía en la actividad por decremento respecto a lo esperado y los valores comprendidos entre 0 y

1 una anomalía por exceso respecto a lo esperado. Un valor de anomalía 0 equivaldrá la identificación de comportamiento normal.

El método de evaluación propuesto, tras realizar la valoración Fuzzy, implementa un proceso de filtrado caracterizado por el parámetro N_L y el parámetro V . Estos dos parámetros definen la forma en la que se realiza la unión y filtrado de anomalías para la detección de tramos temporales de comportamiento anómalo. El proceso de filtrado de anomalías selecciona en la salida solamente aquellos periodos de anomalías consecutivas superiores en longitud a N_L y de valor de anomalía superior a V . El uso de un filtrado de valor permite, en función de la aplicación, filtrar valores bajos de anomalías cuando se desea un sistema de detección más permisivo.

El filtro propuesto permite la identificación de periodos de comportamiento anómalo significativos y podrá ser adaptado en función del caso de uso. Así, el método de detección de anomalías se trata de un método adaptativo al usuario. Por ejemplo, en el caso de un usuario con insomnio, el método de detección de anomalías puede ser parametrizado para ser más restrictivo en la detección de comportamiento anómalo durante la noche (definición de valores bajos de N_L y V durante ese periodo) y menos restrictivo durante el día. El método de filtrado dota de versatilidad al método propuesto para la detección de anomalías y permite la introducción de conocimiento experto en el caso de que sea necesario. El diagrama de bloques de esta fase se muestra en la Figura 6.14.

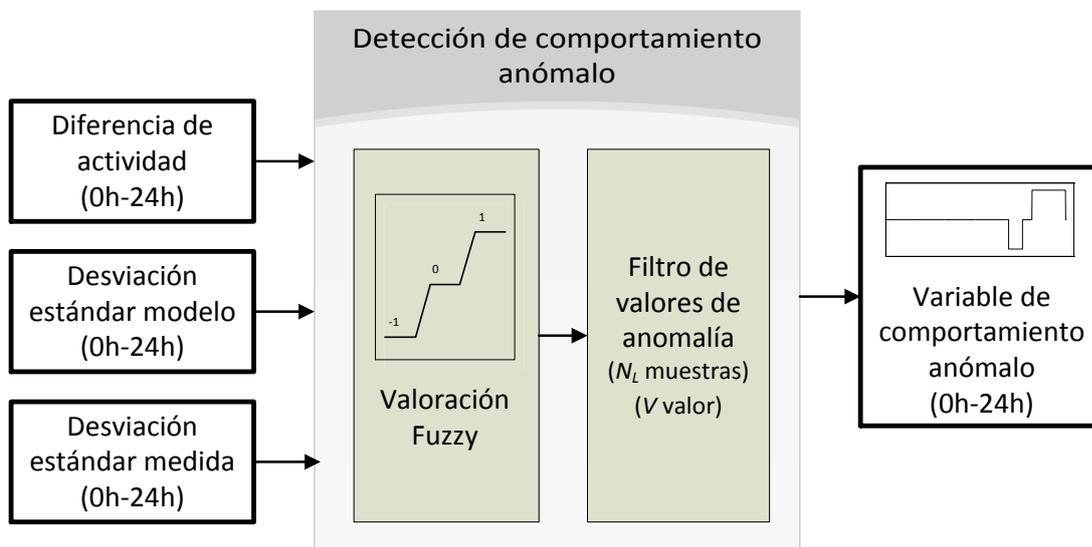


Figura 6.14. Diagrama de la fase III: detección de comportamiento anómalo

```

Function detectarComportamientoAnómalo(Double[1440] DiferenciaActividad,
    Double[1440] SDrutina, Double[1440] SDevaluable, Integer NL = 60, Double V = 1):
    Float x, a, b, c, d = 0;
    Float[1440] ValorAnomalía, ComportamientoAnómalo;
    // Para cada muestra se construye de forma adaptativa la función
    // Fuzzy de valoración
    For Integer i from 1 to 1440:
        x = DiferenciaActividad[i];
        a = -max(SDrutina[i], SDevaluable[i]);
        b = -min(SDrutina[i], SDevaluable[i]);
        c = min(SDrutina[i], SDevaluable[i]);
        d = max(SDrutina[i], SDevaluable[i]);
        ValorAnomalía[i] = valorarAnomalía(x, a, b, c, d); //ver Ecuación 6.6
    End for;
    // A partir de las valoraciones realizadas se realiza el filtrado para identificar
    // comportamientos anómalos (ver Figura 6.14)
    ComportamientoAnómalo = filtrarAnomalías(ValorAnomalía, NL, V );
    Return ComportamientoAnómalo;
End function;

Function filtrarAnomalías(Double[1440] ValorAnomalía, Integer NL = 60, Double V = 1):
    //Filtrar valores menores que V
    For Integer i from 1 to 1440:
        If abs(ValorAnomalía[i]) ≤ V :
            ValorAnomalía[i] = 0; //comportamiento normal
        End if;
    End for;
    //Filtrar segmentos menores que NL muestras
    Integer i, j = 1;
    While (i ≤ 1440):
        If ValorAnomalía[i] ≠ 0:
            j = i;
            While ValorAnomalía[j] ≠ 0:
                j++;
            End while;
            If (j-i) < NL:
                ValorAnomalía[from i to j] = 0;
            End if;
            i = j + 1;
        Else:
            i = i + 1;
        End if;
    End while;
    Return ValorAnomalía;
End function;

```

Figura 6.15. Pseudocódigo de la fase III: detección de comportamiento anómalo

La Figura 6.15 muestra el pseudocódigo para la implementación de la fase completa de detección de comportamiento anómalo. La función *detectarComportamientoAnómalo* recibe la diferencia de actividad calculada en la fase I del método (*DiferenciaActividad*) y las desviaciones estándar calculadas en la fase II (*SD_{rutina}* y *SD_{evaluable}*) junto con los parámetros N_L y V implicados en la detección del comportamiento. Estos dos parámetros son utilizado para filtrar los valores de anomalía en la función *filtrarAnomalías* obteniendo la salida del método: variable de comportamiento anómalo (*ComportamientoAnómalo*). Este proceso de filtrado permite tener en la variable de comportamiento anómalo solo aquellos segmentos detectados cuya duración temporal sea superior a N_L y cuya graduación de comportamiento sea mayor que V .

6.3.3.5. RESUMEN

El método de evaluación propuesto en este trabajo se basa en el cálculo de la diferencia entre la medida de la actividad y la rutina identificada para el usuario (cálculo de residuos). Considerando la rutina como el modelo de comportamiento esperado, este cálculo será similar al cálculo de los residuos del modelo. Las principales limitaciones de este cálculo son resueltas en este trabajo mediante el uso de técnicas de lógica difusa que permiten la valoración del comportamiento de forma automática y adaptativa.

En este trabajo se propone un método semi-supervisado para la evaluación de anomalías que resuelve las principales limitaciones identificadas en el estado del arte en la aplicación de este tipo de métodos. Asumiendo como comportamiento normal a la rutina identificada mediante el método descrito en el Capítulo 5, el método propuesto permite identificar tramos de comportamiento anómalo. Estos tramos pueden ser filtrados en función de las necesidades y del ámbito de trabajo.

En el Apéndice H (sección H.4) se presenta un ejemplo de aplicación del método propuesto sobre un usuario de ejemplo. El pseudocódigo que debe implementarse para la detección de comportamientos anómalos se muestra en la Figura 6.16.

Para un día d de análisis, mediante el establecimiento de los parámetros N_{min} , N_m , N_r , N_L y V se aplican los métodos propuestos en este trabajo. Para ello se obtiene por un lado la actividad sobre la que se desea realizar la detección (*ActividadEvaluable*), se calcula la rutina durante los $N_r = 5$ días anteriores (*Rutina*) y se aplica el método de detección de anomalías propuesto. Las funciones *obtenerActividadFísica* (ver Figura 4.8), *calcularRutina* (ver Figura 5.25), *calcularDiferenciaActividadFiltrada* (ver Figura 6.9), *extraerCaracterísticasEstadísticas* (ver Figura 6.10) y *detectarComportamientoAnómalo* (ver Figura 6.15) descritas en los capítulos anteriores son utilizadas en este pseudocódigo.

```

Día d // Día en el que se desea realizar la detección de comportamientos anómalos
Integer Nmin = 1020; // 70% del día necesario para no eliminar día completo
Integer Nm = 60; // Resolución requerida para la detección
Integer Nr = 5; // Número de días a utilizar para el cálculo de la rutina
Integer NL = 60; // Tamaño mínimo de comportamiento anómalo a detectar
Integer V = 0; // Valor mínimo de anomalía para la detección

// Obtención de la medida de actividad a evaluar
Double[1440] ActividadEvaluable = obtenerActividadFísica(d);

// Cálculo de la rutina
Double[1440] Rutina = calcularRutina(d, Nmin, Nr, Nm);

// Detección de anomalías
// Fase I
Double[1440] DiferenciaActividad =
calcularDiferenciaActividadFiltrada(Rutina, Actividad, Nm);
// Fase II
Double[1440] SDevaluable = extraerCaracterísticasEstadísticas(ActividadEvaluable, Nm);
Double[1440] SDrutina = extraerCaracterísticasEstadísticas(Rutina, Nm);
// Fase III
Double ComportamientoAnómalo = detectarComportamientoAnómalo(DiferenciaActividad,
SDrutina, SDevaluable, NL, V);

```

Figura 6.16. Pseudocódigo para la aplicación del método de detección propuesto

6.3.4. VALIDACIÓN DEL MÉTODO PROPUESTO

El procedimiento seguido para la validación del método de detección de anomalías es similar al seguido en el Capítulo 4 y Capítulo 5 y que se describe en el Apéndice G. Para llevar a cabo la validación del método de análisis propuesto en este trabajo, se ha realizado la monitorización de 12 usuarios (6 hombres y 6 mujeres) pertenecientes al grupo de usuarios adultos según (Welk et al. 2012). Las características de los usuarios se detallan en la sección F.3 del Apéndice F. La Tabla 6.7 muestra el resumen de las características que definen a los usuarios involucrados en el sistema. Cada uno de los usuarios fue monitorizado durante 2 días separados en media 21,75 días ($\sigma = 7,02$).

Tabla 6.7. Usuarios involucrados en la validación del método de detección

	Media	σ
Edad	29,25 años	2,30 años
Peso	67,50 Kg	12,81 Kg
Estatura	1,70 m	0,08 m
Índice de masa corporal	22,71	3,16
Duración del experimento	25,91 horas	3,26 horas
Separación temporal	21,75 días	7,02 días

La monitorización en días separados temporalmente se llevó a cabo para maximizar la posible existencia de cambios de comportamiento y así, analizar en mayor rango el funcionamiento del método propuesto. Para cada día, todos los usuarios fueron monitorizados al menos 24 horas con 25,91 horas en media ($\sigma = 3,26$) mediante un dispositivo vestible con apariencia de reloj y acelerómetro integrado (Texas Instruments 2014). Dentro de este dispositivo se implementó el método descrito en el Capítulo 4 para la medida de la actividad física.

Con el objetivo de evitar en la experimentación errores ocasionados por el modelo de comportamiento generado y focalizar la experimentación en la validación del método de detección, la rutina (modelo de comportamiento) ha sido generada con un valor $N_r = 1$. Así, de los dos días de experimentación, la medida de actividad registrada el primer día es directamente identificada como rutina y la medida del segundo día es la variable sobre la que se detectan comportamientos anómalos (actividad evaluable).

Utilizando los cuestionarios de actividad descritos en el Apéndice G se calcula la diferencia de actividad reportada por el usuario para cada hora. Los tramos temporales para los cuales la diferencia es distinta de cero representan los tramos horarios para los que el usuario realiza un cambio de comportamiento. Adicionalmente, se ha pedido a los usuarios que indiquen en el cuestionario explícitamente los tramos horarios cuyas actividades realizadas el segundo día de monitorización se consideran anómalas respecto al primero.

Puesto que los usuarios solamente pueden indicar el nivel de actividad cada hora, el parámetro N_m toma el valor recomendado de 60 muestras (1 hora). Por esta misma razón, el tamaño del filtro de anomalías es también de $N_L = 60$. El parámetro V se fija con valor nulo ($V = 0$) para que no se lleve ningún filtrado por valor. Así, el filtro de anomalías es configurado para detectar todos los tramos de comportamiento anómalo de duración superior a 1 hora.

Los resultados de este experimento han sido estudiados de acuerdo a la estrategia propuesta en (Ward, Lukowicz, and Gellersen 2011). La métrica de eventos propuesta es utilizada considerando los cambios de comportamiento de los cuestionarios como los cambios esperados que deberían ser detectados por el método propuesto. La variable de comportamiento anómalo detectada por el método será la variable a validar.

Para la aplicación de la métrica propuesta en (Ward et al. 2011) es necesario definir la tolerancia temporal. La tolerancia temporal se define como la variación temporal entre el instante en el que se detecta un comportamiento anómalo y en el que el usuario lo

indicó en los cuestionarios. Puesto que el usuario solamente puede indicar niveles de actividad cada hora, la tolerancia temporal será de 60 muestras. Cada evento (periodo de comportamiento anómalo) detectado por el método de detección de anomalías podrá ser etiquetado como:

- *Evento correcto (C)*. Representa los verdaderos positivos (*true positives, TP*), periodos de comportamiento anómalo detectados por el método y reportados por el usuario.
- *Evento insertado (I)*. Representa los falsos positivos (*false positives, FP*), periodos de comportamiento anómalo detectados por el método y no reportados por el usuario.
- *Evento borrado (D)*. Representa los falsos negativos (*false negatives, FN*), periodos de comportamiento anómalo no detectados por el método y reportados por el usuario.

A parte de la métrica propuesta en (Ward et al. 2011), la validación del método de detección tiene lugar siguiendo dos métricas estándar: precisión y sensibilidad (*recall*). Estas dos métricas, como se mencionó en la sección 2.5.3, permiten validar la fiabilidad del método y pueden ser resumidas de forma conjunta mediante el valor-F. La precisión de un método de detección representa la fracción de eventos que el método es capaz de detectar correctamente (ver Ecuación 6.7). La sensibilidad, también conocida como cobertura, *recall* o exhaustividad, representa la fracción de eventos que el método es capaz de detectar respecto al total que debería de recuperar (ver Ecuación 6.8). El valor-F, también conocido como medida-F o *F-score*, permite mostrar en un mismo valor el efecto de la precisión y la sensibilidad y se calcula, como se muestra en la Ecuación 6.9, a partir de la media armónica de ambas medidas.

$$\text{Precisión} = \frac{\text{Eventos correctos}}{\text{Eventos detectados}}$$

Ecuación 6.7

$$\text{Sensibilidad} = \frac{\text{Eventos correctos}}{\text{Eventos ocurridos}}$$

Ecuación 6.8

$$\text{Valor } F = 2 \cdot \frac{\text{Precisión} \cdot \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}}$$

Ecuación 6.9

La primera fase de validación del método propuesto se basa en la comparación de la salida del método y la diferencia medida en los cuestionarios de actividad. La Tabla 6.8 muestra los resultados obtenidos en la aplicación del método de detección para cada usuario y la Figura 6.17 la representación en base a la métrica seleccionada. El 100% de las salidas proporcionadas por el método son correctas de acuerdo a los cambios indicados en los cuestionarios y solamente dos eventos fueron no detectados (8%). Los valores obtenidos en la experimentación son de una precisión del 100% y una sensibilidad del 92%. Así, el valor-F será 0,96 mostrando un buen compromiso entre alta precisión y alta sensibilidad.

Tabla 6.8. Resultados en base a los cuestionarios de actividad

Usuario	Eventos correctos (TP)	Eventos insertados (FP)	Eventos borrados (FN)
1	3	0	0
2	1	0	0
3	4	0	0
4	1	0	0
5	3	0	0
6	0	0	0
7	1	0	1
8	4	0	0
9	3	0	0
10	1	0	0
11	2	0	0
12	1	0	1
TOTAL	24	0	2

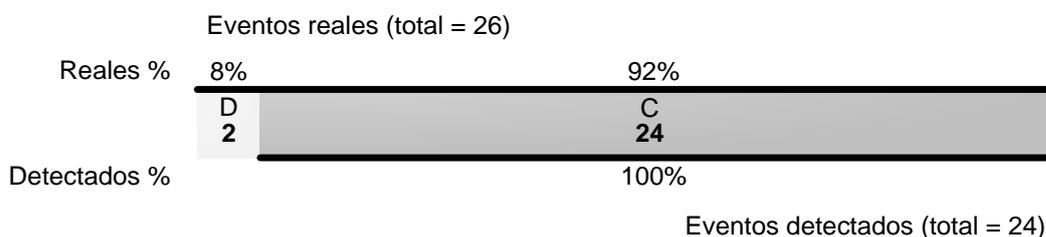


Figura 6.17. Diagrama de análisis de eventos

De acuerdo a la función de valoración Fuzzy descrita en la sección 6.3.3.4 cada uno de los comportamientos anómalos detectados es puntuado con un valor comprendido entre -1 y 1. Un valor 0 corresponde con el comportamiento normal, un valor en el intervalo $[-1, 0)$ con un decremento anómalo de la actividad y un valor en el intervalo $(0, 1]$ con un incremento anómalo de actividad. La relación entre el valor dado por los usuarios y la diferencia proporcionada por los usuarios mediante los cuestionarios fue estudiada con el objetivo de determinar si la función Fuzzy seleccionada da valores adecuados o no. Para ello, se estudia la relación entre el valor dado por la función y la diferencia de actividad reportada por el usuario.

En la Figura 6.18 se representan ambas variables y se aprecia la asociación lineal existente entre ellas que se prueba mediante el cálculo del coeficiente de correlación. Este coeficiente toma un valor de 0,87 ($p\text{-value} < 0,05$) demostrando que existe una fuerte asociación entre ambas variables. De esta forma, se demuestra que la función Fuzzy utilizada en este trabajo para la valoración y basada en una función de pertenencia trapezoidal resulta adecuada para la detección y graduación del comportamiento anómalo.

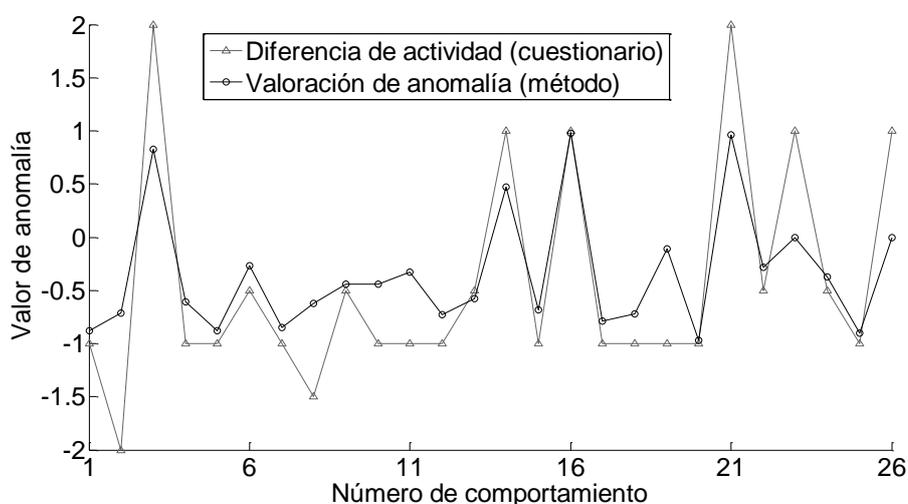


Figura 6.18. Relación entre diferencia de actividad y valor de anomalía

De los 26 cambios de comportamiento detectados a partir del análisis de los cuestionarios, solamente 9 de ellos fueron considerados como verdaderos comportamientos anómalos por los usuarios. Todos estos comportamientos junto con su justificación se muestran en el Apéndice F. Teniendo en cuenta esta información, la precisión real del método propuesto de acuerdo a la percepción de los usuarios sería del 37,5% con una sensibilidad del 100% ($\text{valor-F} = 0,55$). Así, pese a que el método propuesto puede ser considerado como fiable de acuerdo a las medidas tomadas en los

cuestionarios, según la opinión de los usuarios solamente un 37,5% de los comportamientos anómalos detectados se corresponderían con comportamientos anómalos reales. Para resolver este problema y teniendo en cuenta que la resolución de los cuestionarios es horaria, se mantiene en $N_i=60$ el único parámetro que se puede variar en la detección será el parámetro V del proceso de filtrado.

Siguiendo con la métrica propuesta, la Tabla 6.9 muestra los resultados de precisión, sensibilidad y valor-F a medida que el parámetro V toma valores desde 0 (valor inicial) hasta 1 (valor máximo). La detección de eventos para cada usuario se encuentra detallada en la sección F.3 del Apéndice F. Analizando los resultados de precisión y sensibilidad según va variando el parámetro V se construye la gráfica de la Figura 6.19. El parámetro óptimo de trabajo será aquel en el que la curva de precisión y la curva de sensibilidad se crucen. En este caso tiene lugar cuando V es igual a 1, es decir cuando solamente se detectan las anomalías de comportamiento de mayor magnitud. Fijando $V=1$, se consigue tener a la salida del método de detección una precisión y sensibilidad del 100% (valor-F = 1). Los 9 comportamientos identificados por los usuarios como anómalos son detectados con una alta precisión y sensibilidad por el método cuando este se configura con $V=1$.

Tabla 6.9. Resultados del método de detección variando el parámetro V

	$V = 0$	$V = 0,25$	$V = 0,50$	$V = 0,75$	$V = 1$
Precisión	37,5%	45,0%	64,3%	75,0%	100,0%
Sensibilidad	100,0%	100,0%	100,0%	100,0%	100,0%
Valor-F	0,55	0,62	0,78	0,86	1,00

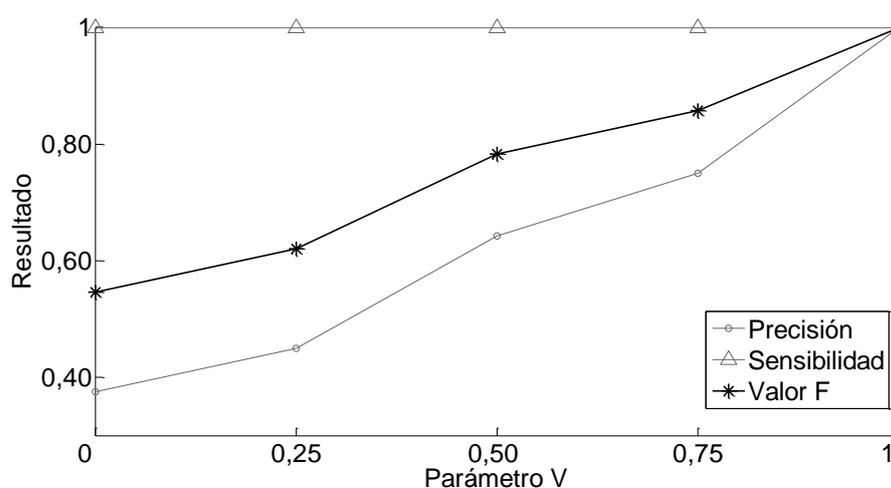


Figura 6.19. Variación del parámetro V en el método de detección

Así, mediante esta etapa de experimentación se puede concluir afirmando que el método de detección de anomalías propuesto permite la detección de todos los cambios de actividad identificados en los cuestionarios (precisión = 100%, sensibilidad = 92%), que existe una fuerte asociación entre la diferencia de actividad dada en los cuestionarios y el valor de anomalía dado por el método (coeficiente de correlación de 0,87, p -value < 0,05) y que la precisión y sensibilidad de acuerdo a los comportamientos anómalos reportados por los usuarios es máxima cuando $V = 1$ en el proceso de filtrado del método de detección.

El método propuesto en este capítulo es capaz de detectar correctamente todos los cambios de actividad que tienen lugar en el comportamiento del usuario con una precisión del 100% y una sensibilidad del 92%. En el caso de que se deseen detectar solamente los comportamientos anómalos, el proceso de filtrado de anomalías debe ser configurado con $V = 1$ consiguiendo una precisión y sensibilidad del 100%.

6.3.5. RESUMEN

La Figura 6.20 muestra el diagrama de bloques del método propuesto dividido en tres fases diferenciadas: cálculo de la diferencia de actividad, extracción de características estadísticas y detección de comportamiento anómalo. Este método recibe como entrada la medida de actividad en la que se desea detectar el comportamiento anómalo y el modelo de comportamiento esperado para el usuario.

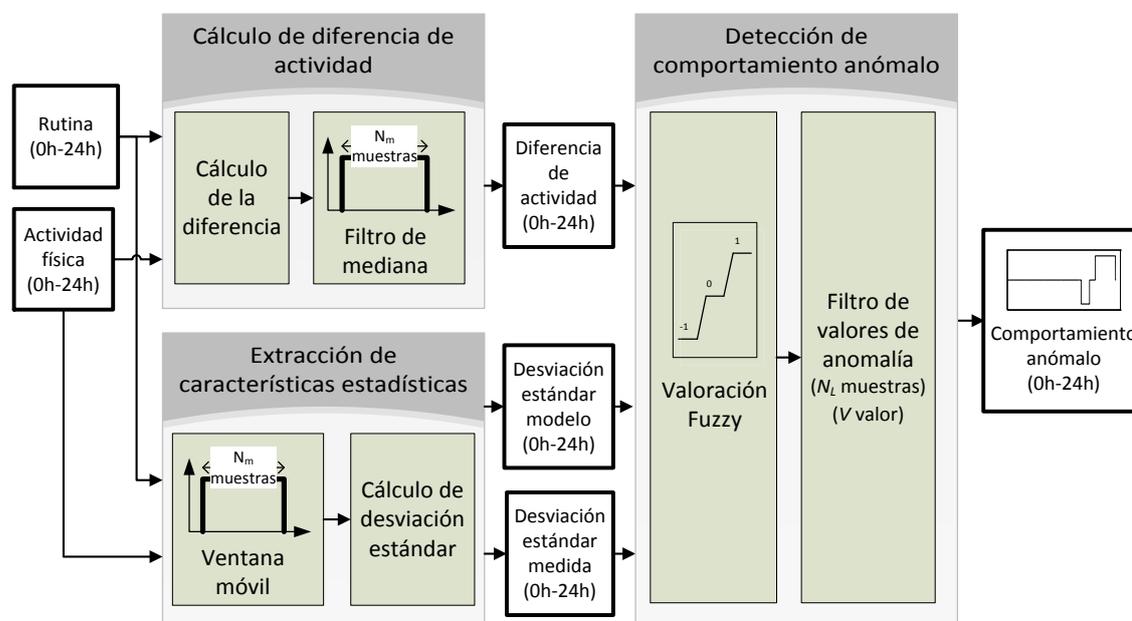


Figura 6.20. Diagrama del método para la detección de anomalías

En el método de detección de anomalías propuesto en este capítulo intervienen tres parámetros que permiten realizar su configuración: N_m , N_L y V . La descripción de estos parámetros y sus valores propuestos se muestran en la Tabla 6.10.

Tabla 6.10. Resumen de los parámetros implicados en el método de detección

Parámetro	Valor propuesto	Descripción
N_m	60	Número de muestras relacionado con la resolución que se desea en el método
N_L	60	Tamaño mínimo de un segmento de anomalías para ser considerado comportamiento anómalo
V	1	Puntuación mínima de anomalía para detectar un comportamiento anómalo

La Tabla 6.11 muestra los requisitos que debe satisfacer un método de análisis del comportamiento humano para la detección de anomalías. Con el método propuesto se satisfacen todos los requisitos identificados en el Capítulo 2.

Tabla 6.11. Requisitos funcionales del método de análisis

Referencia	Requisito	Descripción
<input checked="" type="checkbox"/> RF-7	Método de acondicionamiento	<i>Se implementan las mismas técnicas de pre-procesamiento descrita en la sección 5.3.2.2</i>
<input checked="" type="checkbox"/> RF-8	Método de evaluación	<i>El método propuesto para la detección de anomalías permite evaluar la relación entre el modelo esperado y la medida de actividad física con una alta precisión y una alta sensibilidad. Presenta además una baja tasa de falsas alarmas y una alta tasa de detección</i>
<input checked="" type="checkbox"/> RF-9	Anomalías detectadas	<i>La salida del método propuesto permite identificar instantes temporales cuya actividad medida es anómala y los une para la detección de tramos de comportamiento anómalo</i>

6.4. CONCLUSIONES

El método de detección de anomalías propuesto en este trabajo permite la detección de comportamientos anómalos en la actividad de un usuario. Con comportamiento anómalo se refiere a la detección de niveles de actividad no esperados de acuerdo a la rutina. El método propuesto parte de la medida de actividad física en un día en el que se desean detectar anomalías (*día evaluable*) y el modelo de comportamiento esperado (*rutina*). A través de tres fases con objetivos diferenciados (*cálculo de diferencia de actividad, extracción de características estadísticas y detección de comportamiento anómalo*), el método propuesto es capaz de identificar a la salida los tramos temporales con comportamiento anómalo.

El método propuesto se basa en técnicas estadísticas y de minería de datos aprovechándose de sus ventajas y mejorando sus inconvenientes. Para ello, en la primera fase del método (*cálculo de diferencia de actividad*) se realiza el cálculo y filtrado de los residuos entre el modelo (*rutina*) y la actividad física medida (*actividad evaluable*). La salida de esta fase es evaluada en la última fase del método (*detección de comportamiento anómalo*) mediante una función de valoración Fuzzy construida de forma adaptativa a partir de las características estadísticas extraídas en la fase II del método (*extracción de características estadísticas*) y un proceso de filtrado configurable para cada usuario. La salida del método propuesto es una variable continua con valores comprendidos en el intervalo $[-1, 1]$ y que determina la presencia de comportamiento anómalo:

- *Valor 0*. Se identifica un comportamiento normal.
- *Rango $[-1, 0)$* . Se identifica un comportamiento anómalo relacionado con la detección de un nivel de actividad más bajo que el esperado.
- *Rango $(0, 1]$* . Se identifica un comportamiento anómalo relacionado con la detección de un nivel de actividad más alto que el esperado.

La principal aportación de este trabajo es el diseño de un método de evaluación de anomalías que funciona de forma automática y adaptativa en función de las medidas realizadas. Su principal ventaja respecto al resto de aproximaciones es la creación de una función de valoración Fuzzy cuya construcción depende en cada instante temporal de los valores tanto de la rutina identificada como de la medida de actividad. Esta función Fuzzy se complementa con un proceso de filtrado configurable que permite modificar el comportamiento del método de detección en función de las necesidades particulares que pueda presentar cada usuario.

El método propuesto ha sido validado en un entorno real con 12 usuarios de diferentes perfiles de actividad. En esta etapa de experimentación se ha demostrado su fiabilidad para la detección de cambios de comportamiento con una sensibilidad del 92% y una precisión del 100% (valor-F = 0,96). En la detección de los comportamientos anómalos identificados por los usuarios se ha conseguido una precisión y sensibilidad del 100% (valor-F = 1; $V = 1$ en el filtrado de anomalías). La función de valoración Fuzzy ha sido validada estudiando la relación entre el valor de anomalía dado por el método y la diferencia de anomalía indicada por el usuario mediante cuestionarios (índice de correlación de 0,87; $p\text{-value} < 0,05$).

CAPÍTULO 7.

VALIDACIÓN DEL MÉTODO PROPUESTO

7.1. INTRODUCCIÓN

Partiendo de los métodos de medida, modelado y detección propuestos en el Capítulo 4, Capítulo 5 y Capítulo 6 en este capítulo se realiza la validación global del método propuesto en este trabajo. Todos estos métodos fueron validados en los capítulos en los que fueron presentados. Así, en el Capítulo 4 se validó el uso de JIM como estimador de la actividad física, en el Capítulo 5 se validó el uso del método de modelado propuesto para la identificación de rutinas y en el Capítulo 6 se probó el método propuesto para la detección de anomalías.

El objetivo de este capítulo, tal y como se representa gráficamente en la Figura 7.1 es la validación del método propuesto para la detección de anomalías (métodos de medida, modelado y detección) en base a los requisitos de monitorización identificados en el Capítulo 2: fiabilidad, baja invasividad y alta autonomía. Para ello se trabajará con cuatro personas mayores de 60 años viviendo de forma independiente en su hogar durante un periodo de tiempo prolongado (90 días).



Figura 7.1. Objetivo principal del Capítulo 7

La necesidad de realizar monitorizaciones en escenarios reales y durante periodos de tiempo largos se enuncia en (S Brownsell and Bradley 2011) como uno de los principales objetivos que debe perseguir un sistema de monitorización. Generalmente, las experimentaciones llevadas a cabo en el estado del arte se basan en la monitorización de usuarios en escenarios supervisados y/o durante pocos días/semanas. Tal y como se comenta en (Atallah and Yang 2009) pocos trabajos presentan una experimentación con datos de más de un mes y con usuarios viviendo de forma independiente en su hogar, como es el caso de la validación presentada en este trabajo.

7.2. DISEÑO DEL EXPERIMENTO

7.2.1. ARQUITECTURA DE MONITORIZACIÓN

Los métodos de medida, modelado y detección son validados mediante la monitorización de usuarios en condiciones de vida libre. Para ello, es necesaria la implementación de la arquitectura presentada en la sección 2.4 del Capítulo 2 y que se detalló en (Carús and Díaz 2014).

La Figura 7.2 muestra los elementos que componen el prototipo utilizado en la experimentación. Como elemento de medida se ha seleccionado un dispositivo vestible comercial (TI eZ430-Chronos), un PC actúa como elemento concentrador y la funcionalidad del elemento remoto es implementada a través de una colección de servicios web en la nube.

A cada usuario involucrado en la experimentación se le dio el dispositivo vestible seleccionado y se instaló en su hogar un ordenador portátil que actúa como elemento concentrador.



Figura 7.2. Prototipo de arquitectura para la monitorización

Como elemento de medida se ha seleccionado un dispositivo vestible integrado en un elemento comúnmente vestido por los usuarios: reloj. El dispositivo utilizado (TI eZ430-Chronos) integra diferentes tipos de sensores entre los que destaca el acelerómetro triaxial (CMA3000-D01; rango = ± 8 g) cuyos valores serán monitorizados en este prototipo. El TI eZ430-Chronos posee también una interfaz de comunicación inalámbrica por radio (< 1 GHz RF SoC) y una memoria interna de 8kB para el almacenamiento de medidas. Presenta, por tanto, todas las características necesarias para su uso como elemento de medida y para la implementación del método de medida propuesto. La Tabla 7.1 muestra un resumen de las características técnicas del dispositivo utilizado. El protocolo diseñado para la comunicación con el elemento concentrador es a intervalos regulares (periodo seleccionado de 2 horas) garantizando la integridad y seguridad de los datos.

Tabla 7.1. Resumen de las características técnicas del TI eZ430-Chronos

	Microcontrolador	Arquitectura de programación basada en CC430F6137
	Comunicación	Interfaz radio < 1 GHz SoC (433 MHz)
	Interfaz	Apariencia de reloj deportivo
		Pantalla LCD de 96 segmentos
	Sensores	Sensor de temperatura integrado
		Sensor de batería (voltaje) integrado
		Acelerómetro triaxial VTI CMA3000
		Sensor de presión VTI SCP1000
	Memoria	Memoria flash de 8Kb
Alimentación	Pila de botón CR2032	
Otros	Convertor analógico digital de 8 bits	

El método de medida descrito en el Capítulo 4 es implementado en este dispositivo vestible (ver pseudocódigo en Figura 4.8) para realizar la medida de la actividad física. El acelerómetro es muestreado cada segundo (1 Hz) realizando el cálculo y el almacenamiento de la magnitud propuesta para la medida de la actividad física (JIM) cada minuto. Los requisitos enunciados en la Tabla 4.14 que debe cumplir el acelerómetro para la medida de la actividad física son cumplidos por el acelerómetro utilizado en el dispositivo vestible mencionado. El TI eZ430-Chronos permite el muestreo del acelerómetro a más de 40 Hz y el acelerómetro triaxial que lleva integrado posee un rango de medida mayor que ± 6 g.

El elemento concentrador es implementado en un PC con sistema operativo Windows. Este PC debe estar conectado a Internet y poseer una interfaz de comunicación compatible con el reloj (Figura 7.3.a). Recibe a través de esta interfaz la información monitorizada y la retransmite mediante una comunicación segura al elemento remoto. Un servicio de Windows es ejecutado automáticamente cada vez que el PC es encendido sin que el usuario necesite interactuar con el sistema. De esta forma, cuando el usuario enciende el PC de manera habitual (Figura 7.3.b), la información monitorizada es enviada al elemento remoto de forma transparente para él.

El servicio Windows ha sido desarrollado en C# e instalado en un PC portátil que se ha facilitado a cada usuario (ver Figura 7.3.b). La comunicación con el elemento remoto tiene lugar mediante WiFi en las áreas en las que se tenga acceso y mediante conexión móvil en las áreas que carezcan de conectividad. Dentro del servicio Windows se han implementado técnicas para la detección de fallos y se ha creado una base de datos SQLite que permita el almacenamiento de medidas cuando la conexión con el elemento remoto no esté disponible.



a) Dongle USB para comunicación con el elemento de medida



**b) PC portátil utilizado en las pruebas:
Thinkpad X200 Tablet**

Figura 7.3. Componentes del prototipo de elemento concentrador

El elemento remoto se implementa a través de una colección de servicios web en la nube que almacenan toda la información monitorizada en una base de datos SQL. Este servicio web ha sido implementado utilizando Java para simplificar el despliegue de nuevos servicios web que utilicen la información monitorizada. Así, los métodos propuestos para el modelado y la detección son implementados mediante servicios web Java que acceden a la información de la base de datos.

La plataforma de servicios desarrollada (ver Figura 7.4) consiste en un catálogo de servicios REST que son publicados y accesibles mediante una comunicación HTTPS segura con mecanismos anti CSRF para garantizar la integridad de la información. Los servicios serán consultados por los métodos de procesamiento y análisis para el acceso a la información de actividad. La única forma de acceder a la base de datos en la que se almacena la información será a través de la plataforma de servicios diseñada. De esta forma se garantiza la integridad y protección de los datos almacenados.

El método de modelado y el método de detección son implementados mediante los pseudocódigos de la Figura 5.25 y Figura 6.16. Estos métodos funcionan diariamente en el elemento remoto mediante una tarea programada.

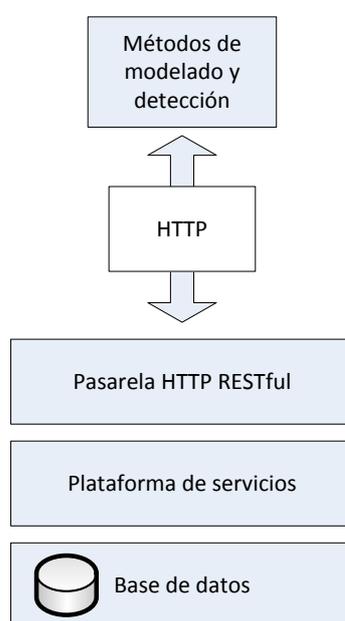


Figura 7.4. Esquema de la arquitectura del elemento remoto

7.2.2. USUARIOS INVOLUCRADOS

La etapa de experimentación tuvo lugar con cuatro usuarios mayores de 60 años durante 90 días. La Tabla 7.2 muestra un resumen de los datos de los usuarios involucrados en el experimento de validación y la Tabla 7.3 la descripción detallada de cada uno de ellos. En el Apéndice I puede encontrarse un resumen de esta información.

Tabla 7.2. Datos resumen de usuarios

Número de usuarios	4
Sexo	3 hombres y 1 mujer
Edad media	69 años
Índice medio de masa corporal	24,53
Duración de la prueba	90 días

Tabla 7.3. Datos detallados de los usuarios

	Usuario 1	Usuario 2	Usuario 3	Usuario 4
Sexo	Hombre	Mujer	Hombre	Hombre
Estado civil	Casado	Casada	Casado	Casado
Profesión	Operario de grúa	Directora de empresa	Administrativo	Operario de maquinaria
Estudios	Sin estudios	Medios	Superiores	Sin estudios
Estado laboral	Jubilado	En activo	Pre-jubilado	Jubilado
Residencia	Ciudad	Ciudad	Ciudad	Ciudad
Edad	72 años	64 años	65 años	78
Estatura	1,60 m	1,72 m	1,76 m	1,65
Peso	75 Kg	54 Kg	77 Kg	70 kg
Índice de masa corporal	29,29	18,25	24,86	25,71

El usuario 1 es un hombre de 77 años, casado y que vive con su mujer. Se encuentra jubilado desde hace más de 10 años y ayuda a su mujer en las tareas del hogar. Se trata de un usuario con sobrepeso al poseer un índice de masa corporal mayor que 25. No tiene un horario claramente predefinido y por la mañana se levanta a diferentes horas dependiendo de cómo haya pasado la noche. En ocasiones se levanta muy pronto ya que periódicamente tiene que ir al médico a hacerse chequeos rutinarios. Todos los días por la mañana acompaña a su mujer a hacer la compra y mientras ella cocina, él limpia la casa y hace el resto de tareas del hogar. Por las tardes, si las condiciones meteorológicas lo permiten sale a dar un paseo con su mujer. En el caso de que no sea posible, pasa la tarde en casa viendo la televisión o durmiendo la siesta. En el área de residencia de este usuario no existe conexión WiFi a Internet y la comunicación entre el elemento concentrador y el elemento remoto tuvo lugar mediante una conexión móvil.

El usuario 2 es una mujer de 63 años, casada y que en la actualidad dirige su propia empresa. Este usuario reside en entorno urbano junto con su marido y su madre. Al

encontrarse laboralmente en activo, presenta una rutina muy pronunciada a lo largo de la semana que solo se ve modificada los fines de semana. Por semana se levanta todos los días a las 9 de la mañana y trabaja en oficina hasta las 7 de la tarde. Cuando sale del trabajo suele quedarse en casa con su madre excepto algunos días que va a clases de yoga. Los fines de semana pasa bastante tiempo fuera de casa, suele quedar con amigos y al menos un día va caminando a una casa que posee en la montaña. Este usuario posee conexión a Internet de banda ancha en su hogar y fue utilizada por el elemento concentrador para la comunicación con el elemento remoto.

El usuario 3 es un hombre de 64 años, casado y que en la actualidad se encuentra prejubilado. Vive en una ciudad junto con su mujer y su suegra. El usuario 3 suele levantarse todos los días alrededor de las 9 de la mañana y tras desayunar, sale a caminar a ritmo fuerte durante más de una hora. A continuación suele hacer ejercicios al aire libre utilizando los aparatos gimnásticos de los que se dispone en los parques públicos de su ciudad. Después de hacer deporte vuelve a casa y sale para encontrarse con amigos hasta la hora de la comida. A veces, ayuda a su suegra a realizar las tareas del hogar. Por la tarde, dependiendo de las condiciones meteorológicas pasa la tarde en casa navegando por Internet o aprovecha para ir a la playa. Antes de cenar, suele dar un paseo por la ciudad. La comunicación entre el elemento concentrador y el elemento remoto tuvo lugar mediante la conexión WiFi que el usuario posee en su hogar.

El usuario 4 tiene 78 años y se encuentra actualmente jubilado. Se despierta todos los días alrededor de las 8 de la mañana y hasta las 10 se queda en casa colaborando con su mujer en las tareas del hogar. A las 10 de la mañana sale a dar un paseo a alta velocidad durante dos horas. A la 1 vuelve de nuevo a casa, come y vuelve a salir alrededor de las 4 o 5 de la tarde para dar otro paseo. Por las tardes camina más despacio y a menor intensidad. Pasa el resto de la tarde en casa viendo la televisión y se acuesta alrededor de las 10 y media de la noche. La comunicación entre el elemento concentrador y el elemento remoto tuvo lugar mediante una conexión de Internet móvil.

7.2.3. PROCEDIMIENTO DE EXPERIMENTACIÓN

El procedimiento utilizado para la experimentación se describe en el Apéndice G. Cada día de monitorización el usuario rellena un cuestionario indicando aproximadamente para cada hora, el nivel de actividad realizado. Los parámetros involucrados en el método propuesto: N_m y N_r son prefijados con los valores ya seleccionados en este trabajo y utilizados principalmente en el estado del arte ($N_m = 60$ y $N_r = 5$). Dentro del

método de detección de anomalías, los parámetros N_L y V son seleccionados en base a los resultados obtenidos en la etapa previa de experimentación (Capítulo 6). Puesto que el usuario indica su nivel de actividad con frecuencia horaria, el parámetro N_L es fijado a 60 muestras. Los cuestionarios de actividad son utilizados para la validación del método propuesto y no se podrán por tanto validar comportamientos de longitud menor a esa. En la etapa de experimentación realizada en la sección 6.3.4 se demostró como todos los comportamientos detectados por el método ($V = 0$) se corresponden con cambios de actividad de acuerdo a los cuestionarios pero para conseguir una máxima fiabilidad en la detección de anomalías de comportamiento el parámetro V ha de tomar su valor máximo ($V = 1$). Así, se fijan $N_L = 60$ y $V = 1$.

La validación llevada a cabo en este capítulo se centrará en los requisitos descritos en el Capítulo 2 y que han guiado el diseño de los métodos propuestos en este trabajo. Se analizará la fiabilidad del método, su invasividad y su autonomía.

El método propuesto para la detección de anomalías se aplica sobre los datos de actividad medidos utilizando los cuestionarios como mecanismo de validación del sistema. Para cada comportamiento anómalo detectado por el método se comprueba su correspondencia con un cambio en el nivel de actividad indicado por el usuario. Esto permite analizar la fiabilidad del sistema a partir del cálculo de su precisión determinando que porcentaje de las anomalías de comportamiento detectadas por el método propuesto se corresponden con comportamientos anómalos reales.

Para validar los requisitos de usuario (no invasividad y autonomía) se utiliza el tiempo de monitorización o tasa de uso del sistema (porcentaje de tiempo que el usuario es monitorizado) y se analizarán los errores detectados. El tiempo de monitorización servirá para determinar la invasividad del sistema analizando principalmente la reluctancia del usuario a ser monitorizado. Los errores detectados permitirán identificar problemas principalmente relacionados con la autonomía del sistema.

7.3. RESULTADOS

7.3.4. FIABILIDAD EN LA DETECCIÓN

La fiabilidad del método propuesto se mide en esta sección a partir del cálculo de la precisión. Esta métrica es calculada a partir del porcentaje de anomalías de comportamiento detectadas por el método que realmente se corresponden con comportamientos anómalos reales. Para ello, cada comportamiento detectado por el

método es cotejado con los niveles de actividad indicados por el usuario en ese tramo temporal para el día en el que se realiza la detección y para los N_r días anteriores.

El Apéndice I muestra todos los cambios de comportamiento detectados para los usuarios siguiendo el formato presentado en la Tabla 7.4. En esta tabla, para cada comportamiento detectado por el método se identifica en la columna “fecha” el número de día de monitorización en el que se detecta el comportamiento y en “día” el día de la semana con el que se corresponde. De esta forma se puede hacer una distinción entre los comportamientos detectados en día laboral (de lunes a viernes) y los detectados en día no laboral (sábado, domingos y días festivos). Cada comportamiento anómalo detectado se identifica a través de sus instantes temporales de detección (“tiempo inicial” y “tiempo final”) y el valor dado por el método en la detección (“valor anomalía”). Puesto que el método ha sido configurado con $V = 1$ este valor siempre será +1 o -1. El nivel de actividad reportado por el usuario en el cuestionario para el tramo temporal en el que se detectó el comportamiento anómalo se muestra en la columna “nivel de actividad día anómalo” y el reportado durante los días anteriores en “nivel de actividad días anteriores”. Este nivel de los días anteriores se calcula como el valor medio dado por el usuario durante los $N_r = 5$ días anteriores de monitorización en el tramo temporal en el que se detectan anomalías de comportamiento. Por último, en la segunda fila de la tabla se muestra la justificación dada por el usuario para ese comportamiento anómalo tras su detección.

Tabla 7.4. Tabla para la presentación de comportamientos anómalos

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores

Partiendo de los parámetros ya indicados en la sección 7.2 ($N_m = 60$, $N_r = 5$, $N_L = 60$, $V = 1$), el método propuesto para la detección de anomalías de comportamiento detecta 56 comportamientos anómalos en los 360 días de monitorización (4 usuarios, 90 días por usuario). Todos los periodos de comportamiento anómalo detectados por el método propuesto han sido validados de acuerdo a los cuestionarios de los usuarios. Para todos los comportamientos detectados como anómalos por el método existe diferencia entre el nivel de actividad reportado por el usuario ese día y el reportado en los días anteriores. Así, la precisión es del 100% ya que todos los comportamientos detectados se corresponden con anomalías de comportamiento reales. Estos cambios anómalos de actividad se corresponden con cambios de rutina que podrían estar relacionados con posibles enfermedades/trastornos. En este caso se ha comprobado que se tratan de

cambios puntuales, tal y como puede verse analizando la justificación de los comportamientos en las tablas del Apéndice I.

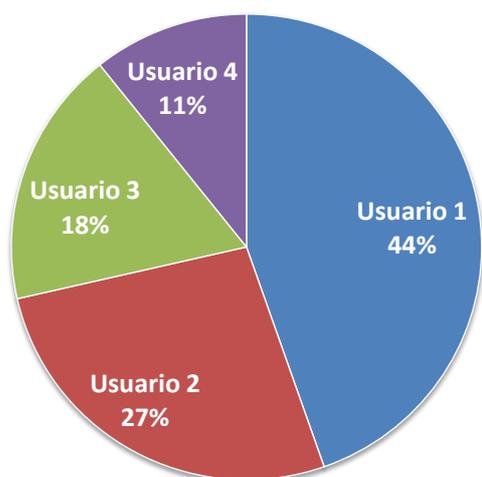
La Tabla 7.5 muestra el número de comportamientos anómalos detectados para cada usuario y su distribución en función del día de la semana diferenciando días laborales (días de semana) y días no laborales (días de fin de semana y días festivos). A su vez, se realiza la diferenciación por tramo horario clasificando de forma independiente las anomalías que tienen lugar en la madrugada (00:00 – 06:00), en la mañana (06:00 – 12:00), en la tarde (12:00 – 18:00) y en la noche (18:00 – 24:00). La Figura 7.5 muestra de forma gráfica la distribución de los comportamientos anómalos detectados por usuario (Figura 7.5.a), por día laboral o no laboral (Figura 7.5.b), por tramo horario (Figura 7.5.c) y el porcentaje de comportamientos validados (Figura 7.5.d).

Tabla 7.5. Comportamientos anómalos detectados y su distribución temporal

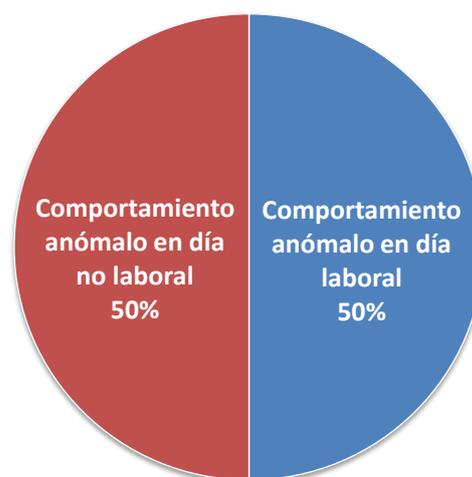
	Total	En día laboral	En día no laboral	En madrugada	En mañana	En tarde	En noche
1	25	76%	24%	0%	52%	28%	20%
2	15	13%	87%	0%	53%	27%	20%
3	10	30%	70%	0%	90%	10%	0%
4	6	67%	33%	0%	50%	50%	0%
	56	50%	50%	0%	59%	27%	14%

A la vista de los resultados mostrados en la Tabla 7.5 y la Figura 7.5, se aprecia como el usuario 1 al poseer una rutina menos marcada es el usuario en el que se detectan un mayor número de comportamientos anómalos. Por el contrario, el usuario 4 es el usuario con el menor número de anomalías de comportamiento detectadas. La distribución de comportamientos anómalos por día muestra que justo el 50% de los comportamientos fueron detectados en día laboral y el 50% en día no laboral. Sin embargo, se aprecia en la Figura 7.5.c que hay un mayor número de anomalías de comportamiento detectadas por la mañana que durante el resto del día. Concretamente, no se ha detectado ningún comportamiento anómalo durante la madrugada. La precisión del 100% en la detección de comportamientos anómalos se visualiza gráficamente en la Figura 7.5.d.

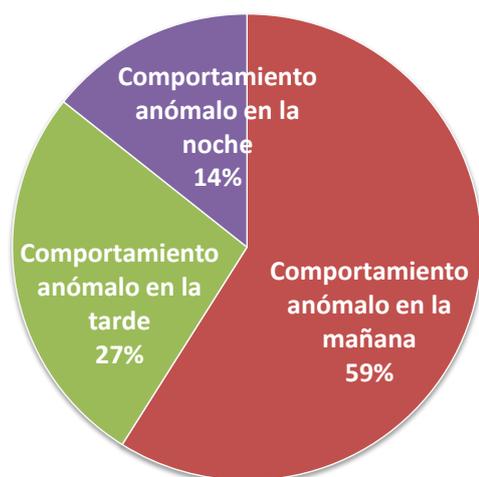
Los perfiles de usuario descritos en la sección 7.2 permiten analizar e interpretar los resultados obtenidos en la validación del método propuesto. Para cada usuario se ha analizado su rutina, el número de comportamientos anómalos detectados y su distribución tanto a nivel de tipo de día (laboral o no laboral) como a nivel de tramo horario (madrugada, mañana, tarde y noche).



a) Distribución de comportamientos anómalos detectados por usuario



b) Distribución de comportamientos anómalos detectados en día laboral y día no laboral



c) Distribución de comportamientos anómalos detectados por tramo horario



d) Distribución de comportamientos anómalos detectados y validados

Figura 7.5. Distribución de los comportamientos anómalos detectados

El usuario 1 fue el usuario con el mayor número de tramos de comportamientos anómalos detectados tal y como se muestra en la Figura 7.5.a. Este usuario no tiene una rutina de comportamiento específica sino que cambia su actividad en función del día y de las condiciones meteorológicas. Esto hace que un gran número de anomalías de comportamiento sean detectadas.

La Figura 7.6 muestra la distribución de comportamientos anómalos detectados, para el usuario 1, por tipo de día (Figura 7.6.a) y por tramo horario (Figura 7.6.b). Los principales cambios de comportamiento detectados en este usuario se producen durante la mañana (08:00 – 12:00) y son principalmente debidos a los cambios en la hora de levantarse ya que varía en función de las tareas que tiene que realizar (ir a la compra, ir al médico...). Por esta razón, el 76% de los comportamientos anómalos tienen lugar durante los días laborales cuando realiza una mayor variación de su rutina.

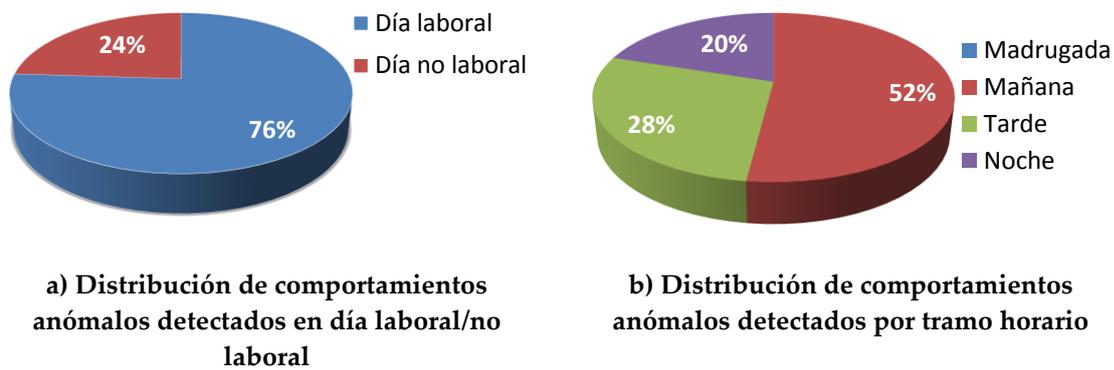


Figura 7.6. Distribución de anomalías detectadas para el usuario 1

El usuario 2 es el único de los usuarios que aún se encuentra activo laboralmente y que tiene, por tanto, una rutina más marcada. Los comportamientos anómalos tienen lugar principalmente durante el fin de semana por la mañana (cuando se levanta tarde) y por las tardes (cuando sale a pasear y a realizar algo de deporte). Así, como se muestra en la Figura 7.7, el 87% de los comportamientos anómalos detectados tienen lugar en día no laboral (fin de semana y días festivos) con un 53% durante el tramo horario de mañana (08:00 – 12:00).

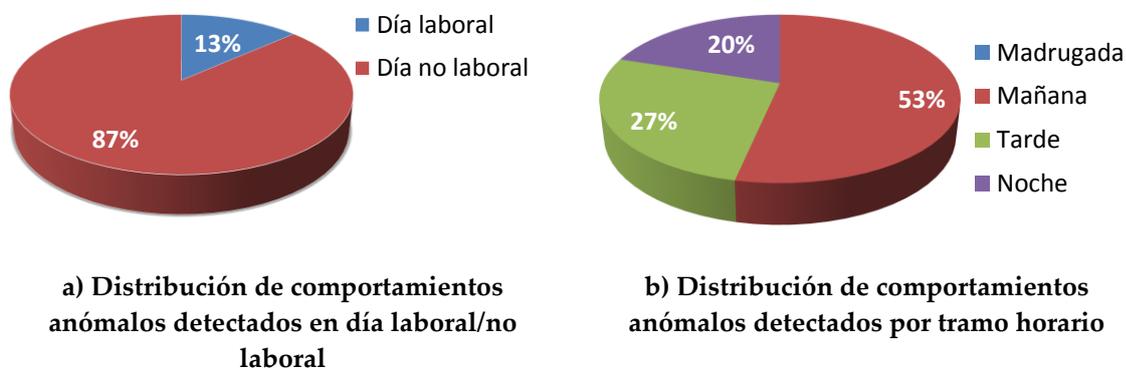


Figura 7.7. Distribución de anomalías detectadas para el usuario 2

El usuario 3 fue uno de los usuarios con el menor número de comportamientos anómalos. Este usuario tiene una rutina muy específica durante las mañanas y la mayoría de los comportamientos anómalos detectados (90%) tienen lugar en este tramo horario cuando modifica su rutina. A su vez, igual que ocurría con el usuario 2, la mayoría de sus comportamientos anómalos tiene lugar durante los días no laborales (70%).

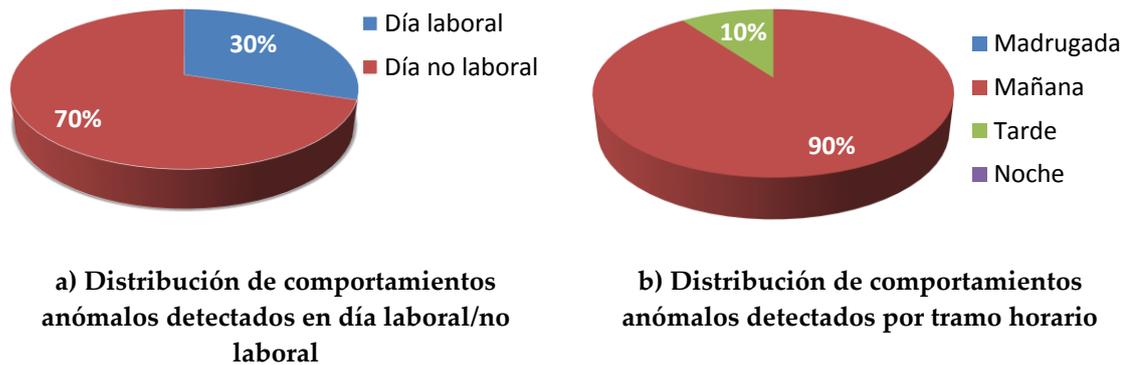


Figura 7.8. Distribución de anomalías detectadas para el usuario 3

El usuario 4 es el usuario con el menor número de comportamientos anómalos. Este usuario sigue con bastante fidelidad su rutina y los comportamientos detectados tienen lugar principalmente durante la mañana y durante la tarde. La Figura 7.9.b muestra la distribución de los comportamientos anómalos detectados por tramo horario: el 50% de los mismos tiene lugar por la mañana y el 50% restante durante la tarde. El usuario modifica levemente su rutina debido a las condiciones meteorológicas que evitan que pueda salir a caminar fuera o que no se sienta en buenas condiciones para hacerlo. Estos cambios de comportamiento son los detectados como anómalos por el sistema y la mayoría tienen lugar durante los días laborales (Figura 7.9.a).

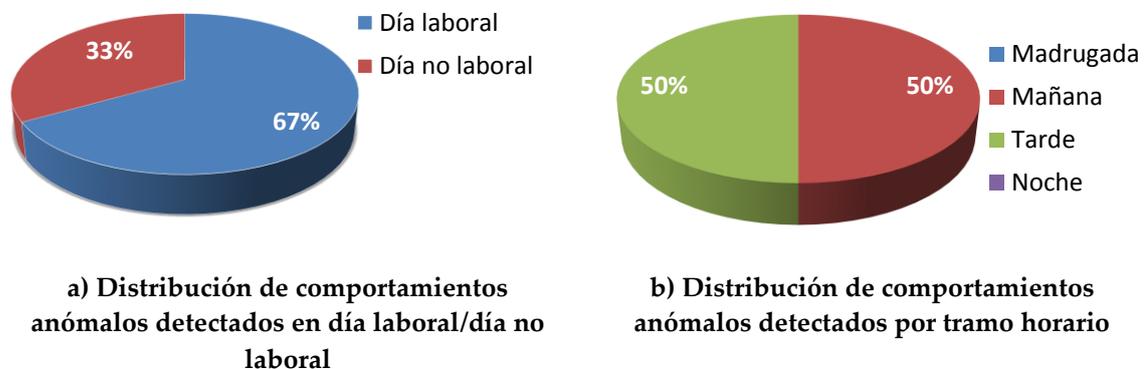


Figura 7.9. Distribución de anomalías detectadas para el usuario 4

Como ya se mencionó, todos los eventos detectados son confirmados como cambios de comportamiento por los cuestionarios por lo que la precisión del método y por tanto su fiabilidad es del 100%. Así, el método propuesto cumple uno de los principales requisitos técnicos establecidos al comienzo del trabajo.

7.3.5. INVASIVIDAD Y AUTONOMÍA EN LA DETECCIÓN

7.3.5.1. INVASIVIDAD

Para medir la invasividad del método se propone el análisis del tiempo de monitorización o tasa de uso del sistema. Este indicador es ampliamente utilizado en el estado del arte para analizar y evaluar el uso de sensores vestibles y en el caso de este trabajo, permite analizar la adherencia y reluctancia de los usuarios a ser monitorizados. La Figura 7.10 muestra una adaptación del algoritmo propuesto en (Hecht et al. 2009) para la detección automática del tiempo de monitorización. Este algoritmo permite determinar para cada uno de los usuarios el porcentaje de tiempo de monitorización. En este trabajo, el tiempo de monitorización es analizado durante las horas de luz y durante el día completo. Se considera el tramo horario 09:00 – 23:00 como horas de luz.

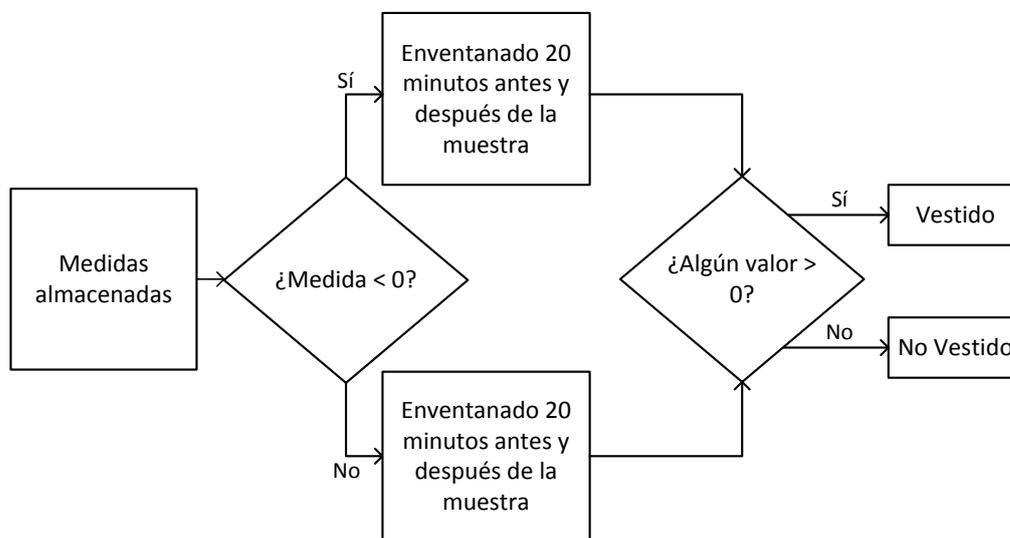


Figura 7.10. Algoritmo para la detección de tiempo de monitorización

Aplicando el algoritmo de la Figura 7.10 sobre los datos de monitorización de los usuarios involucrados en el experimento se construye la Tabla 7.6. Para cada usuario se muestra el porcentaje de tiempo de monitorización durante el día completo y durante las horas de luz. Todos los usuarios fueron monitorizados durante más del 90% de horas de luz con un tiempo medio de monitorización del 63,16% del día completo ($\sigma =$

0,09) y un 95,04% de las horas de luz ($\sigma = 0,08$). Analizando los valores obtenidos, puede verse como la mayor parte del tiempo de no monitorización es debido a las horas de sueño que suponen un porcentaje de alrededor del 40% del día completo (9 horas; 37,5%).

Tabla 7.6. Tiempos de monitorización para cada usuario

Usuario	% Tiempo de monitorización	
	En horas de luz (09:00 – 23:00)	En día completo (00:00 – 24:00)
1	93,46	57,54
2	96,02	65,48
3	93,83	60,61
4	96,84	69,00

El tramo temporal correspondiente con las horas de luz, es el tramo en el que el usuario debería ser monitorizado el 100% del tiempo ya que es cuando se espera que esté despierto. De acuerdo a los valores obtenidos en la monitorización, el usuario quita el reloj un 5% de ese tiempo que se corresponde con alrededor de 45 minutos. Este tiempo se debe principalmente al tiempo en el que usuario se encuentra aseándose. Así, teniendo en cuenta el tiempo en el que usuario se encuentra durmiendo y el tiempo durante el que el usuario se asea, se comprueba el alto porcentaje conseguido de monitorización.

En las Figura 7.11, Figura 7.12, Figura 7.13 y Figura 7.14 se representa para cada día y para cada usuario los porcentajes de monitorización obtenidos en el estudio (tanto durante las horas de luz como durante el día completo).

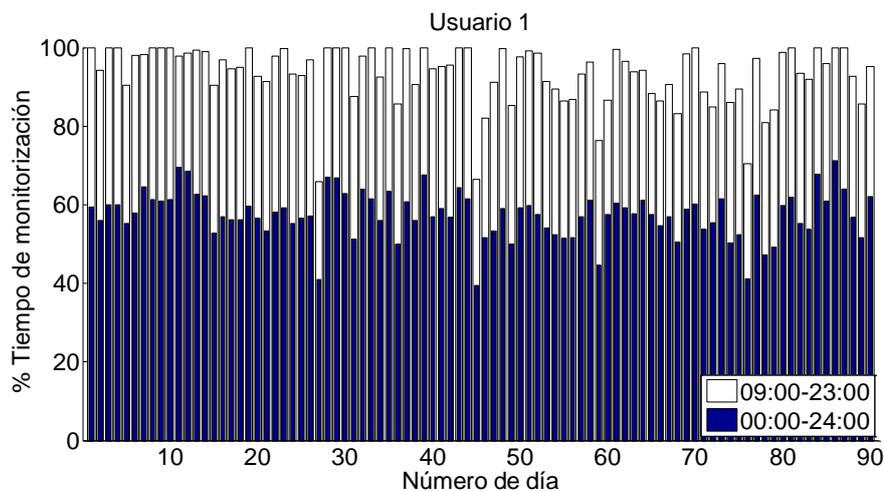


Figura 7.11. Tiempo de monitorización del usuario 1

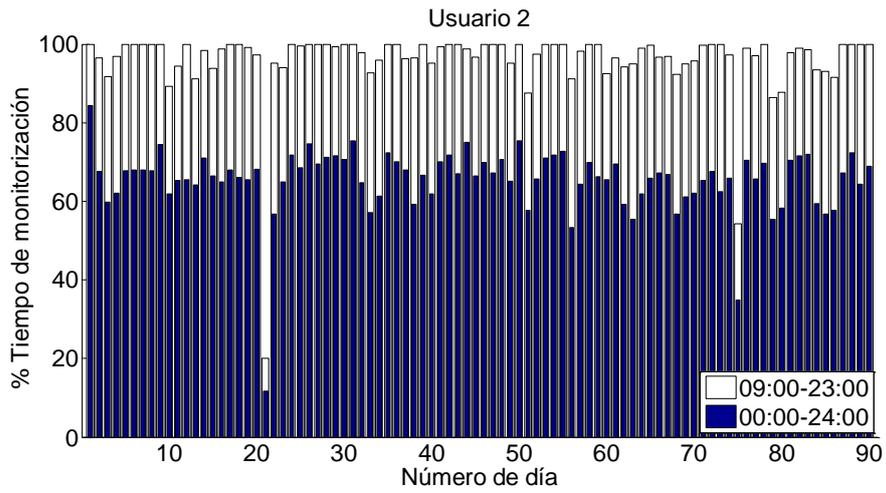


Figura 7.12. Tiempo de monitorización del usuario 2

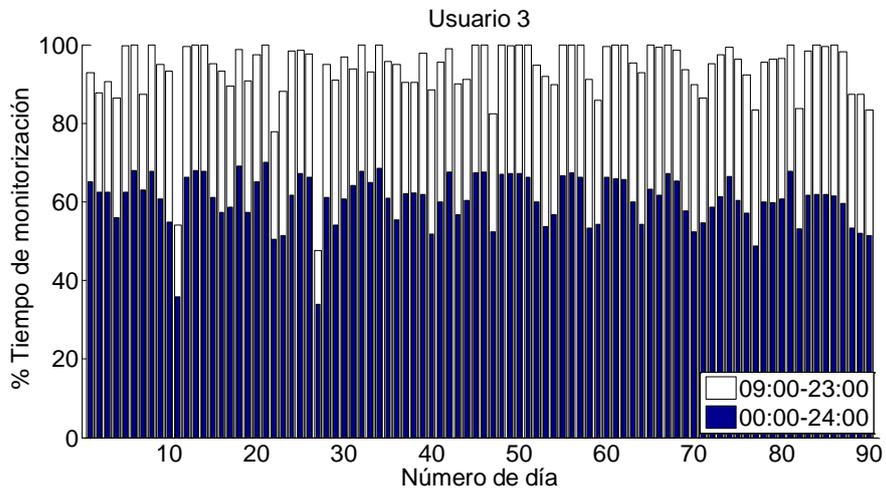


Figura 7.13. Tiempo de monitorización del usuario 3

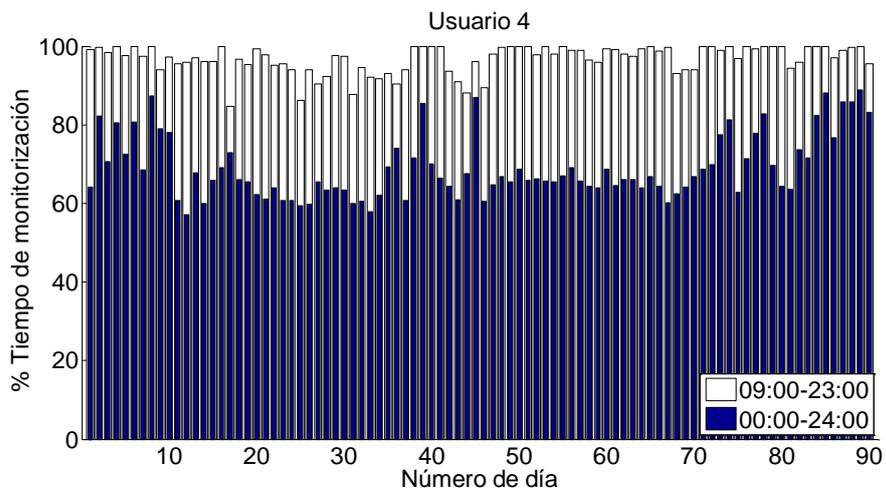


Figura 7.14. Tiempo de monitorización del usuario 4

Dentro de la distribución de tiempos de monitorización, destaca el día número 21 del usuario 2 que presenta valores muy bajos. Se trata de un día en el que el usuario olvidó colocar el dispositivo vestible y que fue detectado dentro de los comportamientos anómalos identificados por el método propuesto.

El estudio realizado en esta sección del tiempo de monitorización muestra como los usuarios son monitorizados durante más del 90% de las horas de luz. Con esto se demuestra la no invasividad del método de monitorización utilizado y la no reluctancia de los usuarios a formar parte de la monitorización.

7.3.5.2. AUTONOMÍA

La autonomía del sistema propuesto es analizada estudiando los errores obtenidos en la monitorización y los días en los cuales el sistema no pudo realizar la monitorización del usuario. Las medidas monitorizadas son recogidas por el elemento concentrador de forma autónoma sin ningún tipo de interacción del usuario. Se identifica la existencia de dos tipos de errores y que ocasionan la existencia de días sin datos:

- *Cambio de pila.* El cambio de pila del elemento de medida debe ser realizado por un usuario experto. Es necesario retirar la tapa trasera del dispositivo vestible utilizando un destornillador para llevar a cabo el cambio de pila. Aunque se han implementado algoritmos para identificar posibles cambios de batería (envío de alarmas en el elemento concentrador), existe un retardo entre la detección y el cambio que provoca que existan días sin datos.
- *Ausencia de sincronización.* Este error es debido a la imposibilidad del elemento vestible de comunicarse con el elemento concentrador (PC del usuario). Debido a la memoria limitada del elemento vestible, cuando el tiempo sin comunicación es prolongado, la monitorización del elemento vestible se detiene. El uso del dispositivo TI eZ430-Chronos permite la monitorización sin envío al elemento concentrador de poco más de 3 días.

La Tabla 7.7 muestra la distribución de los fallos obtenidos durante la validación del método propuesto. El porcentaje de aparición de cada tipo de fallo ha sido calculado respecto al total de días de monitorización (días sin fallos y días con fallos). Para todos los usuarios, los días sin fallos fueron 90 días ya que el experimento finalizó cuando se tenían días suficientes para la validación.

Tabla 7.7. Resumen de fallos de autonomía en la monitorización

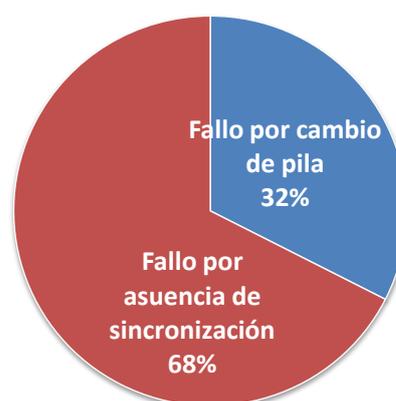
Usuario	Fallos por fin de pila		Fallos por fin de memoria		Total de fallos	
	Nº	%	Nº	%	Nº	%
1	4	4,12	3	3,09	7	7,22
2	3	2,91	10	9,71	13	12,62
3	2	1,98	9	8,91	11	10,89
4	4	4,04	5	5,05	9	9,09
	13	3,25	27	6,75	40	10,00

En total, un 10% de los días de monitorización no pudieron ser registrados debido a fallos relacionados con el fin de pila (cambio de pila) y con el fin de la memoria del dispositivo (ausencia de sincronización). Este último fallo es el que presenta un mayor porcentaje de aparición sobre el total con un 6,75%.

La Figura 7.15.a muestra la distribución de días con fallo en la que se aprecia que un 10% del total de días tuvieron fallo. De esos días en los que hubo un fallo en la autonomía, en la Figura 7.15.b se representa la distribución en función del tipo de fallo ocurrido. Tal y como se mencionaba en el apartado anterior, la mayoría de los fallos ocurridos (68%) están relacionados con la memoria limitada del dispositivo vestible utilizado para la monitorización.



a) Distribución de días con fallo y sin fallo



b) Distribución de tipo de fallo

Figura 7.15. Distribución de fallos de autonomía en la monitorización

Las dos principales razones por las que existen días sin datos están relacionadas con características propias del dispositivo vestible utilizado y no están relacionadas ni con los métodos diseñados ni con la arquitectura de monitorización utilizada. El dispositivo TI eZ430-Chronos seleccionado para la medida de la actividad física permite el almacenamiento de solamente 3 días continuos de datos que aunque cumple con los requisitos establecidos, se comprueba que es un periodo de tiempo bajo para algunos usuarios. Por ejemplo, el usuario 4 suele pasar algunos fines de semana en un pueblo en el que tiene una huerta. Para esos tramos temporales, no se poseen datos de monitorización ya que el dispositivo no es capaz de almacenar los datos de medidas de todos los días que el usuario pasa fuera de su hogar.

Un dispositivo vestible con un sistema de cambio de baterías más sencillo y una capacidad de almacenamiento superior resolvería los fallos detectados en la monitorización. Así, el método y la arquitectura propuesta es autónoma pero otro dispositivo vestible podría mejorar los resultados obtenidos.

7.4. VALIDEZ Y APLICABILIDAD MÉDICA DE LOS RESULTADOS

Los resultados obtenidos en la etapa de experimentación, fueron presentados a una sociedad médica residente en Asturias que ofrece servicios de geriatría y gerontología tanto a pacientes como a sus familiares. De los cuatro usuarios monitorizados, el usuario 1 les pareció el perfil de usuario más interesante desde el punto de vista médico y especialmente como posible afectado por trastornos que podrían ser detectados mediante el método propuesto.

A nivel médico, el usuario 1 es un usuario pre-obeso (índice de masa corporal > 25), con riesgo cardiovascular alto según su historial médico, con un nivel de sedentarismo medio y que padece tres enfermedades crónicas. Es prácticamente independiente aunque presenta un índice de Barthel (Mahoney and Barthel 1965) del 95% a consecuencia de una de las enfermedades crónicas que padece. Este índice de Barthel en el rango [60, 100] le sitúa como un usuario con nivel de dependencia leve y pre-frágil. El síndrome de fragilidad (de Vries et al. 2013) es un concepto que ha surgido en los últimos años como explicación al estado relacionado con el envejecimiento y que se caracteriza por una disminución de la reserva fisiológica. Esta disminución se traduce en un aumento del riesgo de incapacidad del usuario, una pérdida de la resistencia y una mayor vulnerabilidad a eventos adversos como morbilidad o mortalidad.

El método propuesto para la detección de anomalías es identificado por los expertos médicos como una forma de monitorización útil para controlar la actividad del usuario y para detectar cambios de comportamiento significantes. Especialmente en el usuario 1, los expertos médicos consideraron interesante el método propuesto para el análisis de su actividad y la detección de posibles comportamientos anómalos por los siguientes motivos:

- Posee un riesgo cardiovascular alto con episodios críticos en el pasado. Puesto que el sedentarismo es un factor importante en la aparición de enfermedades cardiovasculares, la detección de decrementos anómalos de actividad permite la identificación de un aumento del riesgo de aparición de ataques cardiovasculares en el usuario.
- Según el índice de Barthel, el usuario se encuentra en una situación muy leve de dependencia y es clasificado como pre-frágil (de Vries et al. 2013). La detección de decrementos anómalos de actividad permite, igual que en el caso anterior, la detección de un posible declive funcional, un aumento del riesgo de dependencia del usuario y en consecuencia un decremento de su autonomía.
- Durante el periodo de tiempo en el que se plantea la continuación de la monitorización de actividad del usuario, éste se encuentra inmerso en un tratamiento médico de fuerte impacto relacionado con una de las enfermedades crónicas que padece. Dentro de los efectos secundarios de este tratamiento se encuentra la astenia (cansancio) que conlleva el posible decremento de actividad y los efectos ya mencionados.
- A nivel personal, el usuario se encuentra en una situación de estrés debido al tratamiento médico en el que se encuentra inmerso que le genera ansiedad y susceptibilidad. Padece riesgo de insomnio y de episodios puntuales de depresión. La monitorización de su actividad permite detectar este tipo de comportamientos.

Debido a este interés médico por el uso del método de detección propuesto, se plantea aumentar la duración del experimento durante 90 días más. Así, tras finalizar el experimento cuyos resultados se recogen en la sección 7.3, el usuario 1 continuó siendo monitorizado con el mismo procedimiento descrito en la sección 7.2 durante 90 días adicionales.

El método propuesto fue presentado a los expertos médicos como un método para la detección de comportamientos anómalos de forma adaptativa y automática mediante la monitorización no invasiva y autónoma de usuarios. Los expertos mostraron especial interés por el proceso de filtrado del método de detección de anomalías que

permite su configuración en función de posibles características especiales de los usuarios. Así, mientras que los parámetros N_m y N_r permanecieron iguales que en el resto de experimentos llevados a cabo en este trabajo ($N_m = 60$ y $N_r = 5$), decidieron intervenir sobre los parámetros N_L y V que definen el proceso de filtrado.

Aprovechando las ventajas que presenta el proceso de filtrado y dadas las circunstancias específicas del usuario y la descripción dada en la sección 7.2 sobre su rutina, los expertos médicos determinaron la siguiente configuración (ver resumen en Tabla 7.8):

- *Horario de mañana (8:00 – 16:00)*. Durante este tramo horario el usuario se encuentra realizando las tareas del hogar. Como se vio en los resultados descritos en la sección 7.3, el usuario no siempre comienza su actividad a la misma hora y en ocasiones se detectan pequeños cambios debido a la variabilidad de actividades durante ese tramo horario. Por esta razón, se ha decidido que el filtro sea más permisivo durante este tramo en el que no se consideran que los cambios de comportamiento puedan ser significantes para su estado de salud. N_L será configurado con valor 180 y V con valor 1. Así, solamente se identificarán cambios de comportamiento altamente significantes ($V = 1$) y con una duración superior a 3 horas ($N_L = 180$).
- *Horario de tarde (16:00 – 24:00)*. Durante este tramo horario el usuario realiza actividades de ocio. Generalmente, dependiendo de las condiciones meteorológicas, se queda en su hogar viendo la televisión o sale al exterior a pasear y hacer algo de ejercicio. En este tramo horario la aparición de cambios de comportamiento tiene un alto impacto para su estado de salud. Si el usuario comenzara a pasar más tiempo en su casa, su nivel de sedentarismo aumentaría y en consecuencia su riesgo de dependencia. Debido a esto se establecen los parámetros $N_L = 60$ y $V = 1$. De esta forma se detectarán todos los cambios de comportamiento altamente significantes ($V = 1$) y con una duración superior a 1 hora ($N_L = 60$).
- *Horario de noche (00:00 – 08:00)*. Dadas las circunstancias del usuario y el riesgo de aparición de etapas de insomnio, el filtro ha de ser poco restrictivo durante esta etapa de monitorización para permitir la detección de cualquier cambio de actividad. Para ello se ha configurado el parámetro N_L con valor 60 (1 hora) y el valor V con valor 0. Todos los cambios de comportamiento de duración superior a 1 hora ($N_L = 60$) serán detectados independientemente de su valor de anomalía.

Tabla 7.8. Parámetros del filtro de valores de anomalía

Rango horario	Valor N_L	Valor V	Justificación
00:00-08:00	60	0	Se desean detectar todos los posibles cambios que tengan lugar por las noches por el riesgo de insomnio del usuario
08:00-16:00	180	1	Se desean detectar solamente cambios muy significantes ya que por las mañanas no posee una rutina muy fijada
16:00-24:00	60	1	Se desean detectar todos los posibles cambios de comportamiento que puedan implicar un aumento del riesgo de salud del usuario

Los comportamientos anómalos detectados durante los 90 días adicionales de monitorización son representados en la Tabla 7.9. Con la nueva configuración del proceso de filtrado, se han identificado cuatro comportamientos anómalos: tres de ellos durante la tarde (tramo horario de 16h a 24h) y uno de ellos por la mañana (tramo horario de 8h a 16h). Igual que en los resultados recogidos en la sección 7.3, todos las anomalías de comportamiento detectadas se corresponden con cambios de comportamiento reales y que pueden ser validados (tal y como se muestra en la Tabla 7.9) mediante los cuestionarios de actividad recogidos.

Tabla 7.9. Comportamientos anómalos detectados

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
2	Domingo	16:36	18:18	+1	2,0	0,8
	<i>Llevaba días sin salir a pasear por la tarde debido al mal tiempo</i>					
26	Miércoles	18:36	19:51	-1	1,0	1,8
	<i>Se queda en casa viendo la televisión en lugar de pasear</i>					
47	Lunes	17:21	18:31	-1	0,0	1,7
	<i>Estuvo limpiando las ventanas y quitó el reloj para hacerlo</i>					
89	Viernes	11:12	14:54	+1	0,0	1,8
	<i>No colocó el reloj hasta mediodía en lugar de ponerlo al despertarse</i>					

La configuración realizada por los expertos médicos permite la detección de comportamientos anómalos que están relacionados con los posibles riesgos de salud que presenta el usuario. Así, aunque los cambios de comportamiento son justificados, el método de detección de anomalías propuesto permite detectar cambios susceptibles

de disminuir la calidad de vida del usuario. De acuerdo a las directrices médicas recibidas:

- *Comportamiento anómalo del día 2.* La detección de este comportamiento permite identificar un cambio de comportamiento en el usuario relacionado con el aumento de su actividad física por la tarde. Este cambio es de interés porque permite detectar la adquisición de hábitos saludables por parte del usuario y en consecuencia se conseguiría disminuir el riesgo de aparición de enfermedades.
- *Comportamiento anómalo del día 26.* La detección de este comportamiento permite la identificación de un decremento de actividad en el usuario de acuerdo a su rutina. Este decremento permite la identificación de un aumento del sedentarismo en el usuario. Este cambio tiene lugar debido al mal tiempo que impide que el usuario salga a dar su paseo rutinario al aire libre.
- *Comportamiento anómalo del día 47.* Igual que en el caso anterior se trata de un cambio de comportamiento relacionado con la detección de un decremento de la actividad del usuario. Se trata por tanto de una indicación del aumento del nivel de sedentarismo del usuario.
- *Comportamiento anómalo del día 89.* La detección de este comportamiento permite la detección de un tramo matutino en el que se produce un fuerte y prologando en el tiempo decremento del nivel de actividad del usuario. En este caso, el comportamiento anómalo es debido a que el usuario olvidó colocar el dispositivo de monitorización. Sin embargo, este cambio podría estar relacionado con episodios críticos de alguna de las enfermedades crónicas que padece.

Con la etapa de experimentación realizada y mediante el apoyo médico recibido se ha conseguido validar de nuevo el método de detección de anomalías propuesto. En este caso se ha demostrado su uso y utilidad médica haciendo especial hincapié en la versatilidad y adaptabilidad que introduce en el sistema la configuración del proceso de filtrado. Partiendo de un usuario con varias enfermedades crónicas se ha conseguido detectar comportamientos anómalos susceptibles de modificar sus condiciones de salud.

7.5. DISCUSIÓN

A nivel técnico, en este capítulo se ha demostrado como el método propuesto permite la detección de comportamientos anómalos con una alta tasa de precisión (100%) y por tanto con una alta fiabilidad. Todos los cambios de comportamiento detectados por el método se corresponden con cambios de comportamiento reportados por los usuarios durante el periodo de monitorización (ver Apéndice I). Así, el método propuesto consigue de forma inequívoca la detección de cambios de comportamiento en un usuario que se pueden considerar anómalos.

Para valorar la validez médica del método propuesto se contactó con una asociación de geriatría y gerontología y se le ofreció la oportunidad de utilizar el método para la monitorización de usuarios. Durante 90 días adicionales se monitorizó a un usuario con varias enfermedades crónicas. En este caso el método fue configurado de acuerdo a sus características y a los posibles cambios de comportamiento influyentes en su estado de salud.

Además de los buenos resultados obtenidos, a nivel técnico la principal aportación de este trabajo frente a otros como (Virone 2009) o (Lotfi et al. 2011) es el análisis de datos reales provenientes de usuarios que viven bajo condiciones de vida libre en sus propios hogares. A su vez, se trabaja con datos de varios meses y no del orden de semanas como en la validación de los principales trabajos encontrados en el estado del arte. El método propuesto, a su vez, se centra de forma genérica en la detección de anomalías de comportamiento y no en casos específicos como en (Mubashir et al. 2013), (Taraldsen et al. 2011), (Gil et al. 2013) o (Dias et al. 2012). Su validez y su adaptabilidad para diferentes casos de uso se validó en la sección 7.4.

El tiempo medio de monitorización es una variable analizada en otros trabajos como (Hecht et al. 2009) y (Dias et al. 2012). El último de estos trabajos propone un sistema vestible para la monitorización de pacientes con fibrosis quística en el que obtiene un tiempo de monitorización del 60,4% durante las horas de luz y un 43,8% durante el día completo. Además, 5 de los 15 usuarios analizados durante el experimento rechazaron participar debido a la gran invasividad que suponía el sensor de monitorización seleccionado (colocado en pierna y cadera). En este trabajo, no se ha obtenido el rechazo de ninguno de los usuarios a ser monitorizado y se ha alcanzado una tasa de monitorización superior al 90% en horas de luz y superior al 60% en la medida del día completo.

En (Hecht et al. 2009) el tiempo medio de monitorización obtenido fue del 45% del día completo y el 75% de las horas de luz. En (Troosters, Scieurba, and Battaglia 2010) y

(Waschki et al. 2012) se obtuvieron mayores cifras en el tiempo de monitorización (superiores al 90% del día completo) para pacientes con enfermedad pulmonar obstructiva crónica (COPD). En estos trabajos, el sistema de monitorización ha sido diseñado para ser vestido durante todo el día (incluido mientras duerme). Así, el tiempo de monitorización obtenido será mayor que el que se obtiene en este trabajo pero la comodidad del usuario se reduce considerablemente.

Las principales limitaciones detectadas en la autonomía están relacionadas con el tiempo de vida y la capacidad de memoria del dispositivo utilizado en la monitorización. La necesidad de realizar un cambio de batería en el dispositivo y su limitada memoria ocasionaron respectivamente la pérdida de un 3,25% y un 6,75% de los datos.

En una implementación futura de los métodos propuestos debería considerarse la posibilidad de utilizar técnicas de recolección automática de energía (*energy harvesting*). Existen varias posibilidades incipientes dentro de esta tendencia (Khaligh, Zeng, and Zheng 2010) dentro de las cuales destaca la basada en recolección de energía cinética mediante mecanismos piezoeléctricos (Harb 2011). Otra posibilidad sería el uso de sistemas inductivos de carga de baterías (Hodges 2013). Esta tendencia está comenzando a ser utilizada en la actualidad por muchos sistemas comerciales. Su principal ventaja es que no se necesita ningún tipo de intervención del usuario para el cambio de la batería.

7.6. CONCLUSIONES

Los métodos de medida, modelado y detección propuestos en este trabajo permiten la detección de cambios de comportamiento en personas mayores. Su uso ha sido validado mediante la monitorización de 4 personas mayores durante 90 días en entorno real. Todos los cambios detectados por el método propuesto han sido validados mediante los cuestionarios de actividad cubiertos por los usuarios (ver procedimiento en Apéndice G).

El método propuesto funciona de forma automática y adaptativa detectando cambios de comportamiento sin interacción por parte de los usuarios y con un alto grado de fiabilidad (precisión del 100%). Estos cambios de comportamiento son identificados como anómalos de acuerdo a los niveles de actividad del usuario durante los últimos días. Un cambio en la rutina del usuario no implica una reconfiguración del sistema sino que automáticamente los métodos propuestos se adaptan a las nuevas circunstancias.

La invasividad del método propuesto ha sido analizada estudiando el tiempo durante el cual el usuario ha sido monitorizado. En media, los usuarios fueron monitorizados más del 90% del tiempo en el que se encontraban despiertos. Este valor es superior a trabajos similares existentes en el estado del arte. Se ha comprobado que el uso de sensores integrados en complementos vestibles como relojes, permite minimizar la invasividad de un sistema de monitorización. Así, en este trabajo, utilizando un reloj con apariencia deportiva se ha conseguido una buena adherencia de los usuarios y una muy baja tasa de rechazo.

El único inconveniente identificado en los métodos propuestos está relacionado con el tiempo de vida y la capacidad de memoria del dispositivo seleccionado para la monitorización. Estas características particulares del dispositivo utilizado implicaron la pérdida de un 10% de los datos de monitorización. Por un lado, la duración de la batería es de aproximadamente 2 semanas y se requiere la intervención de un experto para su reemplazo. Un nuevo sistema de alimentación (*energy harvesting* o sistemas de carga inductiva) resolvería este inconveniente. Por otro lado, la capacidad de memoria permite el almacenamiento de solo 3 días de datos. Pese a que esta característica cumple los requisitos identificados, se ha demostrado que es necesaria una mayor memoria que permita a los usuarios mayor flexibilidad. Una mejora en el sistema de alimentación y un aumento en la capacidad de memoria del dispositivo vestible aumentarían la automaticidad del sistema de monitorización.

Uno de los principales avances del método propuesto ha sido el proceso realizado para llevar a cabo su validez médica. Se ha comprobado como el método propuesto puede ser utilizado por expertos médicos para detectar y controlar problemas relacionados con trastornos de comportamiento destacando el síndrome de fragilidad. En la validación llevada a cabo en este capítulo se ha trabajado con un usuario con riesgo cardiovascular alto y con tres enfermedades crónicas. El método fue utilizado exitosamente para identificar situaciones de riesgo del usuario y validado por una asociación médica.

CAPÍTULO 8.

CONCLUSIONES Y TRABAJO

FUTURO

8.1. INTRODUCCIÓN

El envejecimiento progresivo de la población hace necesaria la investigación e inversión en nuevas soluciones basadas en TIC que permitan mejorar la calidad de vida de las personas prolongando su estancia en el hogar y apoyando a los cuidadores en su labor. Las soluciones existentes en la actualidad no están especializadas para su uso en personas mayores y hace que presenten una muy baja tasa de penetración en la sociedad.

La monitorización del comportamiento de un usuario permite realizar de forma eficiente la detección, control y seguimiento de problemas de salud. La detección de anomalías de comportamiento, variaciones de actividad no esperadas, permite intervenir sobre un usuario adelantándose a posibles condiciones adversas y mejorando así su calidad de vida. Una de las posibles magnitudes disponibles para la medida del comportamiento es la actividad física cuya relación bidireccional con el estado de salud de un usuario ha sido ampliamente identificada. A partir de la monitorización de la actividad física se puede llevar a cabo la detección de problemas de salud y mediante la intervención en la actividad física también se puede actuar sobre esos problemas.

Este trabajo propone un método automático y adaptativo para la detección de anomalías de comportamiento a partir de la monitorización de la actividad física

utilizando un sensor de aceleración no invasivo. Los resultados expuestos en el Capítulo 7 muestran la validación de la hipótesis enunciada en la sección 1.3:

Un método de monitorización de comportamiento basado en la medida de la actividad física puede detectar anomalías de comportamiento de forma automática, adaptativa y cumpliendo todos los requisitos necesarios para su uso en personas mayores.

Este capítulo recoge las principales contribuciones de este trabajo y las posibles líneas futuras en las que sería interesante seguir investigando.

8.2. SISTEMAS DE MONITORIZACIÓN DE PERSONAS MAYORES

8.2.1. REQUISITOS IDENTIFICADOS

A partir de un sondeo a más de 80 personas y un minucioso análisis de los trabajos existentes en el estado del arte para la monitorización de personas mayores se han los requisitos necesarios para su aceptación por personas mayores. En el Capítulo 2 de este trabajo se han analizado las ventajas e inconvenientes de todas las aproximaciones encontradas intentando aprovecharse de sus ventajas y proponiendo soluciones que resuelvan los inconvenientes. La colección de requisitos identificados para el diseño de sistemas de monitorización centrados en personas mayores permite el diseño de nuevos e innovadores sistemas de monitorización.

A nivel funcional, se han identificado tres métodos principales de los que debe estar formado un sistema de monitorización: método de medida, método de modelado y método de detección. El método de medida adquiere datos de un sensor y los transforma mediante su procesado en una magnitud de medida que permita monitorizar el estado del usuario. El método de modelado acondiciona y procesa la magnitud adquirida creando un modelo de referencia. Este modelo es utilizado en el método de detección para la evaluación del estado del usuario y para la identificación de posibles condiciones de las que sea necesario notificar o en las que sea necesario intervenir.

Todos los métodos funcionales de un sistema de monitorización (métodos de medida, modelado y detección) han de funcionar de forma automática y adaptativa. Los métodos diseñados deben requerir la mínima interacción en su configuración y

funcionamiento. A su vez, deben adaptarse a las condiciones particulares de cada usuario y ser capaces de auto-configurarse ante posibles cambios en los estados de referencia.

A nivel técnico, los métodos diseñados deben garantizar la seguridad y la fiabilidad. Estos dos aspectos serán claves en la aceptación del sistema tanto por los expertos sanitarios como por los propios usuarios. La baja precisión y sensibilidad de los métodos utilizados en una de las principales debilidades que presentan los métodos actuales.

A nivel de usuario se han identificado ocho requisitos que un sistema de monitorización debe cumplir: comodidad, pequeño tamaño, transparente en su funcionamiento, fácil de utilizar, automático, capacidad de memoria, sistema sencillo de alimentación y amplia cobertura. Estos requisitos pueden resumirse en la búsqueda de un sistema que sea no invasivo para el usuario y que funcione de forma autónoma. Un sistema de monitorización, de acuerdo a los requisitos identificados por los usuarios, debe funcionar de forma transparente sin suponer una molestia en la realización de sus actividades diarias (no invasivo) y debe requerir la mínima interacción en su uso presentando para ello la capacidad de memoria suficiente y un sistema de alimentación adecuado y bien dimensionado (autónomo).

8.2.2. TENDENCIAS EXISTENTES

En el Capítulo 3 se han identificado dos tendencias principales en la monitorización de personas mayores: el sistema conocido como “botón del pánico” y la recolección de variables médicas. Ninguna de estas dos principales tendencias cumple los requisitos identificados y recogidos en la sección anterior:

- El sistema conocido como “botón del pánico” se basa en un dispositivo vestible con un pulsador que el usuario puede utilizar para indicar situaciones de peligro. Cuando una situación crítica es señalizada, el sistema avisa a familiares o a una central de avisos para que actúe ante la situación identificada. Este sistema no funciona de forma automática, no es adaptativo, presenta una alta invasividad y en general poca autonomía.
- El servicio de recolección de variables médicas se basa en una serie de sensores médicos (algunos de ellos vestibles) con los que el usuario toma sus variables de salud. Los valores adquiridos son enviados a expertos médicos de forma transparente para el usuario. Estos expertos pueden consultar de forma telemática las variables del usuario y de forma automática generar avisos o intervenir si es necesario. Este sistema presenta cierto grado de adaptabilidad

pero posee una alta invasividad y aunque funciona de forma semi-automática requiere interacción por parte del usuario.

El análisis del estado del arte en sistemas de monitorización de personas mayores realizado en este trabajo ha permitido identificar las principales debilidades, amenazas, fortalezas y oportunidades de los sistemas y tendencias existentes en la actualidad. La principal debilidad está relacionada con la no adecuación a los requisitos de los usuarios. Esto hace que los sistemas tengan una baja tasa de aceptación y que su principal amenaza sea que los usuarios no sean conscientes de los beneficios que supone su uso. Sin embargo el creciente uso de las TIC y los beneficios reales y demostrados de su uso para la monitorización de personas se convierten en las principales fortalezas poniendo de manifiesto la gran oportunidad que supone su uso.

En este trabajo, a través del análisis de los servicios de monitorización existentes y de los principales trabajos y proyectos de investigación se ha identificado una tendencia incipiente basada en la monitorización y análisis del comportamiento para la detección, control y seguimiento de problemas de salud. Mediante la medida y modelado del comportamiento puede llevarse a cabo la detección de comportamientos anómalos. Con comportamientos anómalos se refiere a la detección de comportamientos no esperados en el usuario de acuerdo a su rutina (modelo de comportamiento) y que pueden estar relacionados con situaciones críticas como caídas, sedentarismo o insomnio.

8.3. MÉTODO DE MEDIDA: ACTIVIDAD FÍSICA

En este trabajo se han identificado los principales sensores y magnitudes utilizados en el estado del arte para la monitorización de personas. Se ha estudiado su aplicabilidad a la medida del comportamiento (objetivo de este trabajo) y se ha valorado su idoneidad para la monitorización de personas mayores de acuerdo a los requisitos identificados.

Dentro de los sensores destacan los sensores de presencia, los sensores de movimiento, los sensores de localización y los sensores de videovigilancia. Estos sensores pueden dividirse en dos categorías principales: sensores ambientales y sensores vestibles. Mientras que los sensores ambientales son los menos invasivos, los sensores vestibles destacan por su precisión y versatilidad. Buscando un compromiso entre ambos, en este trabajo se propone el uso de un sensor vestible que pueda ser integrado en prendas o complementos comúnmente vestidos por el usuario. De esta forma se realiza una monitorización precisa y poco invasiva para el usuario.

La localización, la presencia, el uso de elementos, el consumo eléctrico y la actividad física son algunas de las principales magnitudes utilizadas en el estado del arte para la medida del comportamiento. De todas ellas, destaca la actividad física por tratarse de la única magnitud formal y estándar. Por esta razón, la actividad física es la magnitud seleccionada en este trabajo para la medida del comportamiento. Sin embargo, en el estado del arte esta magnitud ha sido principalmente utilizada con fines deportivos y en estudios descriptivos de la población.

Los métodos para la medida de la actividad física se dividen en métodos estándar, métodos objetivos y métodos subjetivos. Los métodos estándar son los métodos más precisos pero la infraestructura y los procedimientos necesarios para su medida hacen que no sea posible su implementación en entornos reales. Por ello, en este trabajo se han seleccionados los métodos objetivos como los métodos más adecuados para la medida del comportamiento. Dentro de este tipo de métodos se encuentra el uso de sensores fisiológicos y el uso de sensores de movimiento. Buscando un compromiso entre precisión y sencillez y de acuerdo a los requisitos identificados, se propone el uso de acelerómetros para la medida de la actividad física. Sin embargo hay dos limitaciones principales que se identifican en su uso: consumo eléctrico y capacidad de memoria. El sistema de alimentación y la memoria del sensor deben dimensionarse adecuadamente para que el método de medida seleccionado no sea invasivo para el usuario.

Para transformar la aceleración en actividad física es necesaria la existencia de un método de procesado que partiendo de la aceleración medida en el sensor obtenga la medida de actividad. El área de la magnitud de la aceleración (*signal magnitude area, SMA*) es el principal estimador de la actividad física utilizado en el estado del arte y cuya relación con la actividad física ha sido ampliamente validada por la comunidad científica. Sin embargo, este estimador tiene un alto coste computacional al requerir que el acelerómetro sea muestreado a una gran velocidad y que las medidas sean procesadas mediante un doble proceso de filtrado.

En este trabajo se propone un nuevo estimador de la actividad física basado en el cálculo de la derivada de la aceleración (*jerk-based information magnitude, JIM*). Este nuevo estimador presenta un menor coste computacional que el estimador SMA al requerir una tasa de muestreo menor (1 Hz) y al no ser necesario un doble proceso de filtrado. Su validez como estimador de la actividad física se ha demostrado mediante la monitorización de 39 usuarios con diferentes perfiles de actividad en entorno real. La fuerte asociación existente entre el estimador propuesto JIM y el estimador SMA se ha validado mediante el estudio de la correlación (índice de correlación de 0,95; p-value <

0,05), el análisis del sesgo mediante el gráfico Bland-Altman y el análisis de la distribución de valores por usuario. A su vez, el menor coste computacional del estimador JIM se traduce en un menor consumo energético (hasta 5 veces menos, dependiendo del sensor utilizado) y por tanto en un aumento del tiempo de vida del sensor. También, al presentar una menor tasa de muestreo del acelerómetro la capacidad de memoria requerida es optimizada.

Así, el nuevo estimador propuesto en este trabajo (JIM) consigue resolver las dos principales limitaciones identificadas en el uso de acelerómetros: consumo energético y capacidad de memoria.

8.4. MÉTODO DE MODELADO:

IDENTIFICACIÓN DE RUTINA

A partir del uso de la actividad física para la medida del comportamiento, en este trabajo se ha estudiado la identificación de la rutina del usuario. Esta identificación ha de seguir los requisitos identificados al comienzo del trabajo diseñando un método de modelado que permita la identificación automática y adaptativa de la rutina del usuario monitorizado.

El método de modelado ha sido dividido en dos etapas principales: acondicionamiento y procesado. El objetivo de la primera etapa es la transformación de la actividad física para su uso en el método de procesado con el objetivo de garantizar la integridad de los datos con los que trabaja.

En este trabajo se propone un método de acondicionamiento compuesto por una fase de pre-procesamiento, una fase de segmentación y una fase de extracción de características. Para cada una de estas fases se han identificado los principales métodos y técnicas existentes y se ha estudiado su aplicabilidad sobre la medida de la actividad física:

- La fase de pre-procesamiento propuesta se basa en una técnica de limpieza de datos formada por la combinación de una técnica de eliminación de datos redundantes y una técnica de detección y evaluación de datos ausentes. Las técnicas propuestas permiten garantizar la calidad de los datos con los que se trabaja en el modelado y simultáneamente reducir su sesgo. En la fase de segmentación se opta por una técnica de inventariado temporal que permite que el método funcione de forma automática y adaptativa. Para evitar el principal inconveniente de este tipo de técnicas (patrones características

divididos en segmentos diferentes) se propone un proceso de inventariado con solapamiento total de tamaño N_m y ventana cuadrada. La salida de la fase de segmentación será un segmento de tamaño N_m para cada una de las muestras de entrada de la variable de actividad física.

- Dentro de la fase de extracción de características destacan las características temporales debido a su sencillez, fácil interpretación y buen rendimiento. Por este motivo, en este trabajo se propone el uso de la media aritmética como característica a extraer de cada uno de los segmentos obtenidos en la fase de segmentación. Esta característica permitirá conocer la tendencia de la actividad física en cada segmento analizado y puesto que la segmentación se realiza con solapamiento total, la sensibilidad de la media aritmética a la presencia de valores extremos se verá amortiguada. A la salida de la fase de extracción de características se tendrá la media aritmética de las muestras de cada segmento de entrada.

La combinación de la fase de pre-procesamiento, con la fase de segmentación y la fase de extracción de características es equivalente a la aplicación de un filtrado de media sobre la variable original de la actividad física. A la salida de la etapa de acondicionamiento se tendrá para cada muestra de actividad física medida un valor que identifica la tendencia de la actividad medida en su vecindad.

Para el procesado de la actividad física acondicionada y la identificación de la rutina (objetivo del método de modelado) se han identificado en el estado del arte tres tendencias principales: métodos basados en modelado teórico, métodos basados en el reconocimiento de actividad y métodos basados en aprendizaje automático. De todos ellos, los métodos basados en aprendizaje automático son los únicos que pueden funcionar de forma automática y adaptativa. Los métodos basados en modelado teórico implican la incorporación de conocimiento experto en el sistema y los métodos basados en el reconocimiento de actividad requieren una etapa previa de entrenamiento que hace que el sistema no pueda funcionar de forma adaptativa sin re-entrenamiento. A su vez, este tipo de métodos no se adecúan al comportamiento real de un usuario en el que las actividades pueden presentarse de forma concurrente o intermitente y no siempre de manera secuencial.

En este trabajo se propone un método de modelado para la identificación de la rutina del usuario basado en métodos estadísticos. Este método de modelado se basa únicamente en las medidas de actividad física acondicionadas sin realizar ningún asunción sobre la naturaleza de las mismas o sobre su distribución (modelado no paramétrico). A partir de la actividad física medida durante los últimos días (datos

históricos) la rutina del usuario es identificada. El método propuesto se basa en el cálculo de la tendencia de la actividad y construye una variable con el nivel de actividad física esperado para cada instante temporal en función de las observaciones históricas. Esta variable define la rutina del usuario y permite conocer para cada instante temporal el nivel de actividad esperado teniendo en cuenta información “intra-diaria” (medidas a nivel de día) e información “inter-diaria” (medidas en los días anteriores).

El método propuesto en este trabajo para la identificación de la rutina del usuario ha sido validado con 8 usuarios de diferentes perfiles de actividad. En la validación del método propuesto se ha demostrado como de forma automática y adaptativa el método es capaz de identificar la rutina del usuario a partir de las medidas históricas de actividad física. Para llevar a cabo la validación se ha estudiado el índice de correlación entre la rutina identificada y la rutina indicada por el usuario (índice de correlación de 0,92; p -value < 0,05), el análisis de la distribución de valores y el cálculo del índice f de Cohen (valor de 0,88). Siguiendo varias referencias encontradas en el estado del arte, la rutina del usuario es identificada a partir de los datos históricos de $N_r = 5$ días de monitorización.

8.5. MÉTODO DE DETECCIÓN:

ANOMALÍAS DE COMPORTAMIENTO

Con anomalía de comportamiento se hace referencia a la detección de niveles de actividad no esperados según la rutina definida. El método de detección parte del comportamiento medido y del modelo de comportamiento esperado para identificar la existencia de anomalías de comportamiento. Este método es el núcleo principal del método propuesto en este trabajo.

Las principales tendencias identificadas en la detección de anomalías de comportamiento tienen lugar en ámbitos de trabajo muy específicos y con objetivos muy concretos. Existen varios trabajos centrados en la detección de caídas, en la detección de sedentarismo o en la detección de comportamientos asociados a trastornos específicos de insomnio o demencia. Sin embargo, no se han encontrado trabajos que se centren de forma genérica en la detección de anomalías de comportamiento.

Se han identificado tres métodos principales para la detección de anomalías: métodos no supervisados, métodos supervisados y métodos semi-supervisados. Se han

analizado las características de cada uno de estos métodos prestando especial atención a sus ventajas e inconvenientes. Para cada uno de ellos, se han identificado las principales técnicas existentes y su aplicabilidad a la detección de anomalías de comportamiento. Los métodos no supervisados fueron desestimados en la detección ya que se trata de técnicas optimizadas para la agrupación de datos y no para la detección de anomalías como se pretende en este trabajo. Los métodos supervisados son eficientes cuando se trabaja sobre fuentes de datos estáticas con un rango y un dominio claramente definido y no cuando se busca un método automático y adaptativo. En el campo del comportamiento humano estas técnicas son difícilmente aplicables al requerir una etapa exhaustiva de etiquetado y entrenamiento de datos.

Los métodos semi-supervisados han sido seleccionados en este trabajo para la detección de anomalías al permitir la detección automática y no supervisada de las mismas. Este tipo de métodos suponen un compromiso entre los métodos supervisados y los métodos no supervisados asumiendo un modelo de comportamiento normal y detectando como anomalía todas aquellas muestras que difieran de lo esperado. Dentro de los métodos semi-supervisados destacan los métodos estadísticos y cuatro técnicas principales: ajuste de los datos a una distribución conocida, cálculo de histogramas, diagrama de cajas y medidas de similitud. Cada una de estas técnicas ha sido analizada sobre variables de ejemplo de la actividad física estudiando las ventajas e inconvenientes de su aplicación.

La técnica de ajuste de los datos a una distribución normal no es válida para la medida del comportamiento. Se ha demostrado como ni los datos adquiridos ni ninguna de sus más comunes transformaciones pueden ser modeladas por este tipo de distribución. Para ello se ha trabajado con el test Lilliefords que permite evaluar el ajuste de una colección de datos a una distribución de tipo normal. El uso de histogramas para la detección de anomalías ha sido estudiado en este trabajo mediante dos estrategias diferenciadas. Primero se ha estudiado su aplicación sobre el total de medidas de actividad y a continuación sobre segmentos temporales. Pese a que con la segunda de las opciones se han conseguido mejores resultados, el cálculo de histogramas no es una técnica adecuada para la detección de anomalías. Los niveles altos de actividad siempre tienden a presentar una baja tasa de aparición y por tanto, siempre se van a ver puntuados con un valor alto de anomalía. Los diagramas de cajas no presentan tampoco buenos resultados en la detección de anomalías en el comportamiento humano debido a la alta desviación que se encuentra generalmente en las muestras medidas. Dentro de las medidas de similitud destaca la distancia euclídea cuyo valor absoluto es equivalente al cálculo de los residuos de un modelo. El uso de los residuos permite detectar anomalías conociendo el sentido del cambio de nivel de actividad. Sin

embargo, el principal inconveniente de este cálculo es que no se trata de una medida auto-explicativa. Para su uso en la detección de anomalías es necesario la fijación de umbrales u otras técnicas que permitan identificar a partir de qué magnitud de residuo una determinada muestra se debe considerar anómala.

Partiendo de las técnicas analizadas, en este trabajo se propone un método para la detección automática y adaptativa de anomalías dividido en tres fases: cálculo de diferencia de actividad (fase I), extracción de características estadísticas (fase II) y detección de comportamiento anómalo (fase III). En la fase I del método (cálculo de diferencia de actividad), se realiza el cálculo de la diferencia entre las medidas de actividad adquiridas y las esperadas de acuerdo a la rutina identificada para el usuario. Esta diferencia es equivalente al cálculo de los residuos del modelo de comportamiento esperado (rutina). En la fase II se lleva a cabo la extracción de características estadísticas (media y desviación estándar) tanto de la medida adquirida como de la rutina identificada. Estas medidas permiten en la fase III del método (detección de comportamiento anómalo) la creación de una función de evaluación del comportamiento que identifica a su salida los tramos significantes de comportamiento anómalo.

Para el diseño del método propuesto se han realizado varias pruebas de experimentación estudiando y analizando diferentes tipos de funciones de evaluación para a partir del cálculo de residuos identificar tramos de comportamiento anómalo. La función de evaluación propuesta es una función de valoración Fuzzy trapezoidal que se aplica sobre la diferencia calculada en la fase I del método. La función propuesta se construye para cada instante temporal de forma adaptativa en función de las características estadísticas extraídas en la fase II y proporciona un valor entre -1 y 1 proporcional a la anomalía detectada. Un valor 0 identifica normalidad, un valor entre -1 y 0 anomalía por decremento de actividad y un valor entre 0 y 1 anomalía por incremento de actividad.

La función Fuzzy propuesta se complementa con un proceso de filtrado configurable (parámetros N_L y V) que permite el filtrado de anomalías en función de las características propias del usuario y que permite la intervención de expertos en la monitorización. Este proceso de filtrado permitirá la eliminación de tramos de comportamiento anómalo de duración inferior a N_L y cuyo valor de anomalía sea menor que V .

El método de detección propuesto fue validado mediante la monitorización de 12 usuarios adultos con diferentes perfiles de actividad durante dos días. Los cambios de actividad identificados a partir de los cuestionarios de actividad cubiertos por los

usuarios fueron detectados con una precisión del 100% y una sensibilidad del 92% (valor-F = 0,96). Los comportamientos anómalos indicados fueron detectados con una precisión y sensibilidad del 100% configurando el método de detección con $V = 1$ en el proceso de filtrado de anomalías. A su vez, la función de valoración Fuzzy utilizada, basada en una función de pertenencia trapezoidal, fue validada estudiando la asociación entre el valor de anomalía dado por la función y la diferencia de actividad indicada por el usuario en los cuestionarios (índice de correlación de 0,87; p-value < 0,05).

8.6. VALIDACIÓN GLOBAL DEL MÉTODO PROPUESTO

Después de realizar la validación individual de cada uno de los tres métodos propuestos (métodos de medida, modelado y detección), se ha realizado la validación global del sistema utilizando un prototipo de arquitectura para la monitorización de personas mayores. La validación se centró en la monitorización de usuarios en un entorno real y durante un periodo de tiempo prolongado.

La fase de experimentación tuvo lugar mediante la monitorización de cuatro personas mayores con diferentes perfiles de actividad en un entorno real durante 90 días. Todas las anomalías de comportamiento detectadas por los métodos propuestos se correspondieron con cambios de actividad identificados por el usuario y que son explicables teniendo en cuenta su rutina (precisión del método del 100%). Estos cambios han sido detectados de forma automática y adaptativa, requisitos identificados y buscados en la ejecución de este trabajo.

Analizando el porcentaje de tiempo diario en el que el usuario es monitorizado se ha demostrado la no invasividad del método propuesto. En media, los usuarios fueron monitorizados más del 90% del tiempo durante el que se encontraron despiertos, cifra superior a la mayoría de trabajos del estado del arte. El mayor inconveniente identificado durante la monitorización está relacionado con el tiempo de vida y la capacidad de memoria del dispositivo seleccionado en el prototipo de la arquitectura propuesta. Estas dos características hacen que el método posea una tasa de fallo en su funcionamiento del 10%. La duración de la batería es de aproximadamente 2 semanas y se requiere la intervención de un experto para su reemplazo. Una mejora en el sistema de alimentación y un aumento de la capacidad de memoria del dispositivo permitiría aumentar la automaticidad del sistema de monitorización.

Tras la primera fase de experimentación, se presentó el método propuesto a una organización médica que propuso continuar la monitorización con uno de los usuarios

durante 90 días adicionales. Este usuario padece varias enfermedades crónicas y la detección de anomalías de actividad resulta de interés de acuerdo a su historial médico. El método propuesto fue configurado por los expertos sanitarios según las necesidades especiales del usuario monitorizado. Los resultados obtenidos demostraron la validez médica del método al identificar anomalías de riesgo para el usuario monitorizado.

8.7. TRABAJO FUTURO

La etapa de experimentación ha puesto de manifiesto la necesidad de trabajar en un nuevo prototipo de arquitectura. En la actualidad se está trabajando en el diseño de un dispositivo ad-hoc que pueda ser utilizado para la monitorización de personas mayores y que cumpla todos los requisitos identificados en este trabajo. El nuevo dispositivo de medida podrá ser integrado en cualquier complemento de ropa vestido por el usuario y se comunicará con un elemento concentrador en el hogar cuya funcionalidad será básica para actuar únicamente como pasarela de comunicación. Así, el reloj y el ordenador portátil serán sustituidos por elementos hardware más adecuados para la monitorización.

Partiendo de los métodos diseñados, se podría trabajar en el diseño de un servicio de consulta para cuidadores y expertos sanitarios que en tiempo real puedan ver la evolución de las personas monitorizadas. A su vez, una posible línea de trabajo futuro es la investigación en servicios de interés para el propio usuario monitorizado.

Los métodos propuestos han sido diseñados de forma individual de forma que cada uno de ellos podría ser sustituido por nuevos métodos para el desarrollo de nuevos sistemas de monitorización. Dentro de las líneas futuras de investigación está el uso de estos métodos con nuevos objetivos. Por ejemplo, partiendo del método de medida del comportamiento se podría plantear un nuevo método de modelado que identifique el perfil de actividad del usuario (bajo, medio o alto) para en base a él proporcionar servicios de asistencia.

En los próximos meses se prevé la diseminación de los resultados de esta tesis que aún no han sido publicados y la liberación de los datos de experimentación para su uso por la comunidad científica.

8.8. PUBLICACIONES CIENTÍFICAS

Fruto del trabajo de investigación realizado se han desarrollado una serie de publicaciones científicas, tanto en revistas indexadas en el *Journal Citation Report* (JCR) como en congresos.

En el año 2012, en paralelo con el inicio formal de este trabajo se realizó una publicación en la revista "*Computers in Industry*" (JCR = 1.457; Q3) con parte de las técnicas de procesamiento de señales que posteriormente serían estudiadas y analizadas sobre el comportamiento humano. En esta publicación se presenta un método para la valoración de procesos industriales que dio origen a esta tesis:

Molleda, J., Carús, J. L., Usamentiaga, R., García, D. F., Granda, J. C., & Rendueles, J. L. (2012). A fast and robust decision support system for in-line quality assessment of resistance seam welds in the steelmaking industry. *Computers in Industry*, 63(3), 222-230.

Abstract: Assessing the quality of a weld in the steelmaking industry is a complex task. The level of complexity increases when the assessment is based on non-destructive tests. Skilled technicians are often required to make a decision based on automatic assessments of welds. Technicians consider the results of the automatic assessments and use their expert knowledge in order to make a final decision about the quality of the weld. In this paper we propose a decision support system to assess the quality of resistance seam welds of steel strips based on statistical analysis of both the mechanical and electrical variables involved in the welding process to be assessed as well as previously recorded historical data of similar welds. The proposed system is designed following component model based software architecture. The system consists of a set of orthogonal modules: welding variable measurement, welding variable processing and welding quality assessment, communicated by means of dedicated interfaces. The proposed system has been installed in three steel manufacturing lines. With the reduction in the time spent by technicians to make a decision about each weld, the productivity of the manufacturing line has greatly improved. Furthermore, production costs have been reduced since the number of defective welds assessed as non-defective was reduced, and thus the failures in the manufacturing lines due to weld breakages. The experimental results after two years of use in a steel strip galvanizing line are shown.

El estimador propuesto para la medida de la actividad física en la segunda parte del Capítulo 4 fue presentado en Mayo de 2012 en Oporto en la conferencia internacional

pHealth2012: “9th International Conference on Wearable Micro and Nano Technologies for Personalized Health”. Esta publicación fue incluida dentro de un capítulo de libro:

Carús, J. L., Peláez, V., López, G., & Lobato, V. (2012). JIM: a Novel and Efficient Accelerometric Magnitude to Measure Physical Activity. *Studies in health technology and informatics*, 189. IOS Press 2012, 283-288.

Abstract: There are various techniques available to measure human physical activity (PA). Accelerometer based techniques claim to be non-invasive and easy to use. The signal magnitude area (SMA) is the most extended feature used to measure the physical activity. It is calculated by sampling and filtering an accelerometer signal of at least at 50 Hz. SMA has a proven and widely accepted linear relation with the energy expenditure. A novel magnitude called JIM, which is more efficient than SMA, is proposed in this paper. The jerk-based inactivity magnitude (JIM) is also calculated from the acceleration signal, but at a sampling rate of 1Hz, increasing the battery life of the measuring system. This magnitude gives the same information as the SMA (correlation of 95%) and is validated with a group of 39 users in free-living conditions for at least 24 hours.

En la siguiente edición del mismo congreso pHealth2013: “10th International Conference on Wearable Micro and Nano Technologies for Personalized Health” el método de detección de anomalías fue presentado como continuación del trabajo. En este artículo se presentó una primera versión del método propuesto en la segunda parte del Capítulo 6. Esta publicación fue incluida, tal y como se indica a continuación, como capítulo de libro:

Carús, J. L., Peláez, V., López, G., Fernández, M. A., Álvarez, E. & Díaz, G. (2013). Unsupervised Abnormal Human Behaviour Detection Using Acceleration Data. *Studies in health technology and informatics*, 189. IOS Press 2013, 65-70.

Abstract: Abnormal human behavior detection under free-living conditions is a reliable technique to detect activity disorders and diseases. This work proposes an acceleration-based algorithm to detect abnormal behavior as an abnormal increase or decrease in physical activity (PA). The algorithm is based on statistical features of human physical activity. Using a period of observed physical activity as a reference, the algorithm is able to detect abnormal behavior in other periods of time. The approach is unsupervised as the modeling of the reference behavior is not required. It has been validated with a group of 12 users under free-living conditions for two days. Results show a precision greater than 75% and a recall of 92%.

A continuación, se llevó a cabo la expresión formal de la arquitectura utilizada para la experimentación y cuya implementación se describió en la sección 7.2.1 de Capítulo 7. Esta publicación tuvo lugar en el congreso “*IEEE 10th International Conference on Ubiquitous Intelligence & Computing*”:

Carús, J. L., Peláez, V., García, S., Fernández, M. A., Díaz, G., & Álvarez, E. (2013). A Non-invasive and Autonomous Physical Activity Measurement System for the Elderly. *Proceedings of 2013 IEEE 10th International Conference on Ubiquitous Intelligence & Computing and 2013 IEEE 10th International Conference on Autonomic & Trusted Computing*, 619-624.

Abstract: Invasiveness and lack of autonomy are reported as the main barriers of current physical activity measurement systems. A non-invasive and autonomous physical activity measurement system is proposed in this work. The system is based on the use of a wearable device, a data gateway and a remote service to collect and process data. Using a watch-like device, the proposed solution is validated with three elderly users living independently at home under free-living conditions. The compliance and acceptance of the system is better than similar works.

Como resumen de la investigación realizada en el marco de esta tesis, se ha realizado en Octubre de 2014 la siguiente publicación de impacto en la revista “*Pervasive and Mobile Computing*” (JCR = 1,667; Q2):

Carús Candás, J. L., Peláez, V., López, G., Fernández, M. Á., Álvarez, E., & Díaz, G. (2014). An automatic data mining method to detect abnormal human behaviour using physical activity measurements. *Pervasive and Mobile Computing*, 15, 228-241.

Abstract: Abnormal human behaviour detection under free-living conditions is a reliable method to detect disorders and diseases in healthcare applications. The problem with current methods to detect human behaviour changes is the use of supervised techniques that require human intervention. This work proposes an automatic data mining method based on physical activity measurements. Abnormal human behaviour is detected as an increase or decrease of the physical activity according to the historical data. Human behaviour is evaluated in real time grading its abnormality. The method has been validated involving users with a precision of 100% and a recall of 92%.

Las publicaciones anteriormente descritas se complementan con las siguientes comunicaciones en el ámbito de técnicas de procesamiento de variables, desarrollo de algoritmos y monitorización de usuarios:

- Carús, J. L., García, S., García, R., Waterworth, J., & Erdt, S. (2014). The ELF@ Home project: Elderly sELF-care based on sELF-check of health conditions and sELF-fitness at home. *Studies in health technology and informatics*, 200. IOS Press 2014, 164-166.
- Fernández, M. Á., Peláez, V., López, G., Carús, J. L., & Lobato, V. (2012). Multimodal interfaces for the smart home: findings in the process from architectural design to user evaluation. *Ubiquitous Computing and Ambient Intelligence, Lecture Notes in Computer Science*, 173-180.
- Carús, J. L., Alvarez, I., Fernández, I., Gutiérrez, J. P., Royo, L. J., & Goyache, F. (2006). Diversidad filogenética como método de utilidad en programas de conservación de recursos genéticos ganaderos. *ITEA, información técnica económica agraria: revista de la Asociación Interprofesional para el Desarrollo Agrario (AIDA)* 2, 133-138.

A partir del trabajo desarrollado en esta tesis, CTIC realizó en Noviembre de 2014 la solicitud de una patente admitida a trámite (referencia de solicitud 201400932) de título “Sistema para la monitorización de la actividad física en personas mayores” que nace como continuación del trabajo descrito en este documento y en la que se reconoce al autor de esta tesis como inventor.

REFERENCIAS

- AAL JP. 2015. "Ambient Assisted Living Joint Programme: ICT for Ageing Well." Retrieved February 10, 2015 (<http://www.aal-europe.eu/>).
- Aanesen, Margrethe, Ann Therese Lotherington, and Frank Olsen. 2011. "Smarter Elder Care? A Cost-Effectiveness Analysis of Implementing Technology in Elder Care." *Health Informatics Journal* 17(3):161–72.
- Abbate, Stefano et al. 2012. "A Smartphone-Based Fall Detection System." *Pervasive and Mobile Computing* 8(6):883–99.
- Aggarwal, J. K., and M. S. Ryoo. 2011. "Human Activity Analysis: A Review." *ACM Computing Surveys* 43(3):1–43.
- Ainsworth, B. E. et al. 2000. "Compendium of Physical Activities: An Update of Activity Codes and MET Intensities." *Medicine and science in sports and exercise* 32(9 Suppl):S498–504.
- Ainsworth, B. E. et al. 2011. "Compendium of Physical Activities: A Second Update of Codes and MET Values." *Medicine and Science in Sports and Exercise* 43(8):1575–81.
- Ajzen, I. 1991. "The Theory of Planned Behavior." *Organizational behavior and human decision processes* 50:179–211.
- Alwan, Majd, Devon Wiley, and Jeremy Nobel. 2007. *State of Technology in Aging Services*.
- Analog Devices. 2013. "Hoja de Características Del Acelerómetro ADXL345." Retrieved April 20, 2014 (<http://www.analog.com/en/mems-sensors/mems-inertial-sensors/adxl345/products/product.html>).

- Antonakaki, Panagiota, Dimitrios Kosmopoulos, and SJ Perantonis. 2009. "Detecting Abnormal Human Behaviour Using Multiple Cameras." *Signal Processing* 89(9):1723–38.
- Anuradha, G., and Bidisha Roy. 2014. "Suggested Techniques for Clustering and Mining of Data Streams." Pp. 265–70 in *IEEE 2014 International Conference on Circuits, Systems, Communication and Information Technology applications CSCITA*.
- Armitage, C. J., and M. Conner. 2001. "Efficacy of the Theory of Planned Behaviour: A Meta-Analytic Review." *British journal of social psychology* 40(4):471–99.
- Atallah, Louis, Benny Lo, Rachel King, and Guang-Zhong Yang. 2010. "Sensor Placement for Activity Detection Using Wearable Accelerometers." Pp. 24–29 in *International Conference on Body Sensor Networks*. Ieee.
- Atallah, Louis, and Guang-Zhong Yang. 2009. "The Use of Pervasive Sensing for Behaviour Profiling — a Survey." *Pervasive and Mobile Computing* 5(5):447–64.
- Avci, Akin, and Stephan Bosch. 2010. "Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey." Pp. 1–10 in *23rd International Conference on Architecture of Computing Systems (ARCS)*.
- Aztiria, Asier, Alberto Izaguirre, and Juan Carlos Augusto. 2010. "Learning Patterns in Ambient Intelligence Environments: A Survey." *Artificial Intelligence Review* 34(1):35–51.
- Barnes, N., S. Webster, and T. Mizutani. 2006. "Liverpool Telecare Pilot: Case Studies." *Informatics in Primary Care* 14(3):197–202.
- Berlin, JE, KL Storti, and JS Brach. 2006. "Using Activity Monitors to Measure Physical Activity in Free-Living Conditions." *Physical Therapy* 86(8):1137–45.
- Bernabei, Michel et al. 2010. "Automatic Detection of Tic Activity in the Tourette Syndrome." Pp. 422–25 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*.
- Biswas, Jit et al. 2010. "Health and Wellness Monitoring through Wearable and Ambient Sensors: Exemplars from Home-Based Care of Elderly with Mild Dementia." *Annals of Telecommunications - Annales Des Télécommunications* 65(9-10):505–21.
- Boissy, Patrick, Stéphane Choquette, Mathieu Hamel, and Norbert Noury. 2007. "User-Based Motion Sensing and Fuzzy Logic for Automated Fall Detection in Older

- Adults." *Telemedicine journal and e-health : the official journal of the American Telemedicine Association* 13(6):683–93.
- Bourouis, Abderrahim, Mohamed Feham, and Abdelhamid Bouchachia. 2011. "Ubiquitous Mobile Health Monitoring System for Elderly (UMHMSE)." *International Journal of Computer Science and Information Technology* 3(3):74–82.
- Bouten, C. V., K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen. 1997. "A Triaxial Accelerometer and Portable Data Processing Unit for the Assessment of Daily Physical Activity." *IEEE transactions on bio-medical engineering* 44(3):136–47.
- Brignell, Michael, Richard Wootton, and Len Gray. 2007. "The Application of Telemedicine to Geriatric Medicine." *Age and ageing* 36(4):369–74.
- Brownsell, S., and D. Bradley. 2011. "A Systematic Review of Lifestyle Monitoring Technologies." *Journal of Telemedicine and Telecare* 17(4):185–89.
- Brownsell, Simon, and David Bradley. 2011. "Developing a Systems and Informatics Based Approach to Lifestyle Monitoring within eHealth: Part I - Technology and Data Management." *2011 1st IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology (HISB)* 26–29.
- Busmann, Johannes B. J., Ulrich W. Ebner-Priemer, and Jochen Fahrenberg. 2009. "Ambulatory Activity Monitoring." *European Psychologist* 14(2):142–52.
- Butte, NF, Ulf Ekelund, and KR Westerterp. 2012. "Assessing Physical Activity Using Wearable Monitors: Measures of Physical Activity." *Medicine & Science in Sports & Exercise* 44(1 Suppl 1):S5–12.
- Carús, Juan Luis, and Gabriel Díaz. 2014. "Trabajo de Fin de Máster En Investigación En Ingeniería Eléctrica Electrónica Y Control Industrial: Arquitectura Para La Monitorización No Invasiva Y Autónoma de Personas Mayores."
- Carús, Juan Luis, Víctor Peláez, Gloria López, and Vanesa Lobato. 2012. "JIM: A Novel and Efficient Accelerometric Magnitude to Measure Physical Activity." Pp. 283–88 in *Studies in health technology and informatics*, vol. 177.
- Caspersen, C. J., K. E. Powell, and G. M. Christenson. 1985. "Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research." *Public health reports* 100(2):126–31.
- Catellier, D. J. et al. 2005. "Imputation of Missing Data When Measuring Physical Activity by Accelerometry." *Medicine and science in sports and exercise* 37(11 Suppl):S555–62.

- Chakraborty, Debejyo, and N. Kovvali. 2013. "Generalized Normal Window for Digital Signal Processing." Pp. 6083–87 in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- Chan, Marie, Eric Campo, Daniel Estève, and Jean-Yves Fourniols. 2009. "Smart Homes - Current Features and Future Perspectives." *Maturitas* 64(2):90–97.
- Chandola, V., A. Banerjee, and V. Kumar. 2009. "Anomaly Detection: A Survey." *ACM Computing Surveys (CSUR)* 41(3):15.
- Charlon, Yoann, Nicolas Fourty, and Eric Campo. 2013. "A Telemetry System Embedded in Clothes for Indoor Localization and Elderly Health Monitoring." *Sensors* 13(9):11728–49.
- Chen, C., DJ Cook, and AS Crandall. 2013. "The User Side of Sustainability: Modeling Behavior and Energy Usage in the Home." *Pervasive and Mobile Computing* 9(1):161–75.
- Chen, Chih-Ming. 2011. "Web-Based Remote Human Pulse Monitoring System with Intelligent Data Analysis for Home Health Care." *Expert Systems with Applications* 38(3):2011–19.
- Chen, Kong Y., and David R. Bassett. 2005. "The Technology of Accelerometry-Based Activity Monitors: Current and Future." *Medicine & Science in Sports & Exercise* 37(Supplement):S490–500.
- Chen, Liming, and Ismail Khalil. 2011. "Activity Recognition: Approaches, Practices and Trends." Pp. 1–31 in *Activity Recognition in Pervasive Intelligent Environments*.
- Chikhaoui, Belkacem, Shengrui Wang, and H el ene Pigot. 2012. "ADR-SPLDA: Activity Discovery and Recognition by Combining Sequential Patterns and Latent Dirichlet Allocation." *Pervasive and Mobile Computing* 8(6):845–62.
- Chung, Pau-Choo, and Chin-De Liu. 2008. "A Daily Behavior Enabled Hidden Markov Model for Human Behavior Understanding." *Pattern Recognition* 41(5):1572–80.
- Courtin, Emilie, Nadia Jemai, and Elias Mossialos. 2014. "Mapping Support Policies for Informal Carers across the European Union." *Health policy* 118(1):84–94.
- Coye, Molly Joel, Ateret Haselkorn, and Steven DeMello. 2009. "Remote Patient Management: Technology-Enabled Innovation and Evolving Business Models for Chronic Disease Care." *Health affairs (Project Hope)* 28(1):126–35.
- Craig, Cora L. et al. 2003. "International Physical Activity Questionnaire: 12-Country Reliability and Validity." *Medicine and science in sports and exercise* 35(8):1381–95.

- Cuddihy, Paul, Jenny Weisenberg, Catherine Graichen, and Meena Ganesh. 2007. "Algorithm to Automatically Detect Abnormally Long Periods of Inactivity in a Home." P. 89 in *1st ACM SIGMOBILE international workshop on Systems and networking support for healthcare and assisted living environments - HealthNet '07*. ACM Press.
- Curone, Davide, Gian Mario Bertolotti, Andrea Cristiani, Emanuele Lindo Secco, and Giovanni Magenes. 2010. "A Real-Time and Self-Calibrating Algorithm Based on Triaxial Accelerometer Signals for the Detection of Human Posture and Activity." *IEEE transactions on information technology in biomedicine* 14(4):1098–1105.
- Darkins, A., P. Ryan, and R. Kobb. 2008. "Coordination/home Telehealth: The Systematic Implementation of Health Informatics, Home Telehealth, and Disease Management to Support the Care of Veteran Patients." *Telemedicine and e-Health* 14(10):1118–26.
- Davis, Jonathan J., and Andrew J. Clark. 2011. "Data Preprocessing for Anomaly Based Network Intrusion Detection: A Review." *Computers & Security* 30(6-7):353–75.
- Dean, R. B., and W. J. Dixon. 1951. "Simplified Statistics for Small Numbers of Observations." *Analytical Chemistry* 23(4):636–38.
- Dewitte, K., and C. Fierens. 2002. "Application of the Bland–Altman Plot for Interpretation of Method-Comparison Studies: A Critical Investigation of Its Practice." *Clinical chemistry* 48(5):799–801.
- Dias, André et al. 2012. "Assessing Physical Activity in the Daily Life of Cystic Fibrosis Patients." *Pervasive and Mobile Computing* 8(6):837–44.
- Doukas, Charalampos, and Ilias Maglogiannis. 2008. "Advanced Patient or Elder Fall Detection Based on Movement and Sound Data." Pp. 103–7 in *2008 2nd International Conference on Pervasive Computing Technologies for Healthcare*. Ieee.
- Economist Intelligence Unit. 2011. *The Future of Healthcare in Europe*.
- Ellingsen, Kristian, Grahame J. Coleman, Vonne Lund, and Cecilie M. Mejdell. 2014. "Using Qualitative Behaviour Assessment to Explore the Link between Stockperson Behaviour and Dairy Calf Behaviour." *Applied Animal Behaviour Science* 153:10–17.
- Englert, Heike et al. 2010. "Is Patient Self-Report an Adequate Tool for Monitoring Cardiovascular Conditions in Patients with Hypercholesterolemia?" *Journal of public health* 32(3):387–94.

- Ermes, M., J. Pärkka, J. Mantyjarvi, and I. Korhonen. 2008. "Detection of Daily Activities and Sports with Wearable Sensors in Controlled and Uncontrolled Conditions." *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society* 12(1):20–26.
- Esling, Philippe, and Carlos Agon. 2012. "Time-Series Data Mining." *ACM Computing Surveys* 45(1):1–34.
- Estevez-Tapiador, Juan M., Pedro Garcia-Teodoro, and Jesus E. Diaz-Verdejo. 2004. "Anomaly Detection Methods in Wired Networks: A Survey and Taxonomy." *Computer Communications* 27(16):1569–84.
- European Commission. 2013. "7th European Framework Programme (FP7)." Retrieved June 17, 2014 (http://cordis.europa.eu/fp7/home_es.html).
- European Commission. 2014. "Programa Marco Para La Competitividad Y La Innovación (CIP)." Retrieved June 17, 2014 (http://ec.europa.eu/cip/index_es.htm).
- European Commission. 2015a. "Horizon 2020: The EU Framework for Research and Innovation." Retrieved February 10, 2015 (<http://ec.europa.eu/programmes/horizon2020/en>).
- European Commission. 2015b. "Joint Research Centre (JRC). Institute for Prospective Technological Studies." Retrieved June 17, 2014 (<http://ec.europa.eu/jrc/en/institutes/ipts>).
- Fessl, Angela, V. Rivera-Pelayo, and L. Müller. 2011. "Motivation and User Acceptance of Using Physiological Data to Support Individual Reflection." *2nd MATEL Workshop at European Conference for Technology Enhanced Learning (ECTEL 2011)*.
- Figo, Davide, Pedro C. Diniz, Diogo R. Ferreira, and João M. P. Cardoso. 2010. "Preprocessing Techniques for Context Recognition from Accelerometer Data." *Personal and Ubiquitous Computing* 14(7):645–62.
- Fletcher, Richard Ribon et al. 2010. "iCalm: Wearable Sensor and Network Architecture for Wirelessly Communicating and Logging Autonomic Activity." *IEEE transactions on information technology in biomedicine* 14(2):215–23.
- Fleury, Anthony, Michel Vacher, and Norbert Noury. 2010. "SVM-Based Multimodal Classification of Activities of Daily Living in Health Smart Homes: Sensors, Algorithms, and First Experimental Results." *IEEE transactions on information technology in biomedicine* 14(2):274–83.

- Fontecha, Jesús, and Ramón Hervás. 2011. "A Proposal for Elderly Frailty Detection by Using Accelerometer-Enabled Smartphones." in *5th International Symposium of Ubiquitous Computing and Place Intelligence*.
- Fouquet, Y., and C. Franco. 2010. "Telemonitoring of the Elderly at Home: Real-Time Pervasive Follow-up of Daily Routine, Automatic Detection of Outliers and Drifts." Pp. 121–39 in *Smart Home Systems*. INTECH Open Access Publisher.
- Freedman, David, and Persi Diaconis. 1981. "On the Histogram as a Density estimator:L 2 Theory." *Probability theory and related fields* 57(4):453–76.
- Freescale. 2012. "Hoja de Características Del Acelerómetro MMA8450Q." Retrieved April 20, 2014 (http://cache.freescale.com/files/sensors/doc/data_sheet/MMA8450Q.pdf).
- Fundación CTIC. 2010. *ATLAS - Aplicación de Tecnologías Asistenciales: Informe Final*.
- Gil, Enrique a, Xavier L. Aubert, Els I. S. Møst, and Domien G. M. Beersma. 2013. "Human Circadian Phase Estimation from Signals Collected in Ambulatory Conditions Using an Autoregressive Model." *Journal of biological rhythms* 28(2):152–63.
- Godfrey, A., R. Conway, D. Meagher, and G. OLaighin. 2008. "Direct Measurement of Human Movement by Accelerometry." *Medical engineering & physics* 30(10):1364–86.
- Gogoi, P., D. K. Bhattacharyya, B. Borah, and J. K. Kalita. 2011. "A Survey of Outlier Detection Methods in Network Anomaly Identification." *The Computer Journal* 54(4):570–88.
- Gokalp, Hulya, and Malcolm Clarke. 2013. "Monitoring Activities of Daily Living of the Elderly and the Potential for Its Use in Telecare and Telehealth: A Review." *Telemedicine journal and e-health* 19(12):910–23.
- Grubbs, FE. 1969. "Procedures for Detecting Outlying Observations in Samples." *Technometrics* 11:1–21.
- Gupta, Abhinav, Aniruddha Kembhavi, and Larry S. Davis. 2009. "Observing Human-Object Interactions: Using Spatial and Functional Compatibility for Recognition." *IEEE transactions on pattern analysis and machine intelligence* 31(10):1775–89.
- Györfbíró, Norbert, Ákos Fábíán, and Gergely Hományi. 2008. "An Activity Recognition System For Mobile Phones." *Mobile Networks and Applications* 14(1):82–91.

- Hajj Chehade, Nabil, Pinar Ozisik, James Gomez, Fabio Ramos, and Greg Pottie. 2012. "Detecting Stumbles with a Single Accelerometer." Pp. 6681–86 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Hamm, Jihun, Benjamin Stone, Mikhail Belkin, and Simon Dennis. 2013. "Automatic Annotation of Daily Activity from Smartphone-Based Multisensory Streams." Pp. 328–42 in *Mobile Computing, Applications, and Services*, vol. 110.
- Harb, Adnan. 2011. "Energy Harvesting: State-of-the-Art." *Renewable Energy* 36(10):2641–54.
- Haugen, H. a., L. N. Chan, and F. Li. 2007. "Indirect Calorimetry: A Practical Guide for Clinicians." *Nutrition in Clinical Practice* 22(4):377–88.
- Hautala, Arto J. et al. 2010. "Physical Activity and Heart Rate Variability Measured Simultaneously during Waking Hours." *American journal of physiology. Heart and circulatory physiology* 298(3):H874–80.
- Heart, Tsipi, and Efrat Kalderon. 2013. "Older Adults: Are They Ready to Adopt Health-Related ICT?" *International journal of medical informatics* 82(11):e209–31.
- Hecht, Ariel, Shuyi Ma, Janos Porszasz, and Richard Casaburi. 2009. "Methodology for Using Long-Term Accelerometry Monitoring to Describe Daily Activity Patterns in COPD." *Copd* 6(2):121–29.
- Heil, Daniel P., Soren Brage, and Megan P. Rothney. 2012. "Modeling Physical Activity Outcomes from Wearable Monitors." *Medicine and science in sports and exercise* 44(1 Suppl 1):S50–60.
- Heise, David, and Marjorie Skubic. 2010. "Monitoring Pulse and Respiration with a Non-Invasive Hydraulic Bed Sensor." Pp. 2119–23 in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*. Ieee.
- Helaoui, Rim, Mathias Niepert, and Heiner Stuckenschmidt. 2011. "Recognizing Interleaved and Concurrent Activities Using Qualitative and Quantitative Temporal Relationships." *Pervasive and Mobile Computing* 7(6):660–70.
- Heric, Dušan, and Damjan Zazula. 2007. "Combined Edge Detection Using Wavelet Transform and Signal Registration." *Image and Vision Computing* 25(5):652–62.
- Hodge, VJ, and Jim Austin. 2004. "A Survey of Outlier Detection Methodologies." *Artificial Intelligence Review* 22:85–126.
- Hodges, Steve. 2013. "Batteries Not Included: Powering the Ubiquitous Computing Dream." *Computer* 46(4):90–93.

- Holzinger, A., and G. Searle. 2010. "Perceived Usefulness among Elderly People: Experiences and Lessons Learned during the Evaluation of a Wrist Device." Pp. 1–5 in *4th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*.
- Hong, Yu-Jin, Ig-Jae Kim, Sang Chul Ahn, and Hyoung-Gon Kim. 2008. "Activity Recognition Using Wearable Sensors for Elder Care." Pp. 302–5 in *Second International Conference on Future Generation Communication and Networking*. Ieee.
- Hong, Yu-Jin, Ig-Jae Kim, Sang Chul Ahn, and Hyoung-Gon Kim. 2010. "Mobile Health Monitoring System Based on Activity Recognition Using Accelerometer." *Simulation Modelling Practice and Theory* 18(4):446–55.
- Huo, Hongwei, Youzhi Xu, Hairong Yan, Saad Mubeen, and Hongke Zhang. 2009. "An Elderly Health Care System Using Wireless Sensor Networks at Home." Pp. 158–63 in *2009 Third International Conference on Sensor Technologies and Applications*. Ieee.
- Iglesias, JA, and P. Angelov. 2010. "Human Activity Recognition Based on Evolving Fuzzy Systems." *International Journal of Neural Systems* 20(5):355–64.
- Ikonomovska, Elena, Kurt Driessens, Saso Dzeroski, and Joao Gama. 2011. "Adaptive Windowing for Online Learning from Multiple Inter-Related Data Streams." Pp. 697–704 in *2011 IEEE 11th International Conference on Data Mining Workshops*. Ieee.
- Ilea, Dana E., and Paul F. Whelan. 2011. "Image Segmentation Based on the Integration of Colour–texture descriptors—A Review." *Pattern Recognition* 44(10-11):2479–2501.
- Intille, Stephen S. et al. 2011. "Design of a Wearable Physical Activity Monitoring System Using Mobile Phones and Accelerometers." Pp. 3636–39 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Jakkula, V., and D. J. Cook. 2008. "Anomaly Detection Using Temporal Data Mining in a Smart Home Environment." *Methods of information in medicine* 47(1):70–75.
- Joe, Jonathan, and George Demiris. 2013. "Older Adults and Mobile Phones for Health: A Review." *Journal of biomedical informatics* 46(5):947–54.
- Karantonis, Dean M., Michael R. Narayanan, Merryn Mathie, Nigel H. Lovell, and Branko G. Celler. 2006. "Implementation of a Real-Time Human Movement Classifier Using a Triaxial Accelerometer for Ambulatory Monitoring." *IEEE transactions on information technology in biomedicine* 10(1):156–67.

- Karunanithi, Mohanraj. 2007. "Monitoring Technology for the Elderly Patient." *Expert review of medical devices* 4(2):267–77.
- Kasteren, T. L. M., G. Englebienne, and B. J. a. Kröse. 2010. "An Activity Monitoring System for Elderly Care Using Generative and Discriminative Models." *Personal and Ubiquitous Computing* 14(6):489–98.
- Kayacan, Erdal, Baris Ulutas, and Okyay Kaynak. 2010. "Grey System Theory-Based Models in Time Series Prediction." *Expert Systems with Applications* 37(2):1784–89.
- Keegan, Jennifer et al. 2013. "Improved Respiratory Efficiency of 3D Late Gadolinium Enhancement Imaging Using the Continuously Adaptive Windowing Strategy (CLAWS)." *Magnetic resonance in medicine* 1074(71):1064–74.
- Khaligh, Alireza, Peng Zeng, and Cong Zheng. 2010. "Kinetic Energy Harvesting Using Piezoelectric and Electromagnetic technologies—State of the Art." *IEEE Transactions on Industrial Electronics* 57(3):850–60.
- Khalil, Ashraf, and Suha Glal. 2009. "StepUp: A Step Counter Mobile Application to Promote Healthy Lifestyle." Pp. 1–5 in *2009 International Conference on the Current Trends in Information Technology (CTIT)*. Ieee.
- Khan, Adil Mehmood, Young-Koo Lee, Sungyoung Y. Lee, and Tae-Seong Kim. 2010. "A Triaxial Accelerometer-Based Physical-Activity Recognition via Augmented-Signal Features and a Hierarchical Recognizer." *IEEE transactions on information technology in biomedicine* 14(5):1166–72.
- Khan, SS, ME Karg, Jesse Hoey, and D. Kulic. 2012. "Towards the Detection of Unusual Temporal Events during Activities Using HMMs." Pp. 1075–84 in *2012 ACM Conference on Ubiquitous Computing (UbiComp12)*.
- Khan, Wajahat Ali, Maqbool Hussain, Muhammad Afzal, Muhammad Bilal Amin, and Sungyoung Lee. 2012. "Healthcare Standards Based Sensory Data Exchange for Home Healthcare Monitoring System." Pp. 1274–77 in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*.
- Kim, E., S. Helal, and D. Cook. 2010. "Human Activity Recognition and Pattern Discovery." *IEEE Pervasive Computing* 9(1):1–14.
- King, AC, and DK King. 2010. "Physical Activity for an Aging Population." *Public Health Reviews* (2107-6952) 32(2):401–26.
- Kivel, Antti. 2013. "What Determines Our Health?" in *10th International Conference on Wearable Micro and Nano Technologies for Personalized Health (pHealth2013)*.

- Knudson, Duane V. 2013. *Qualitative Diagnosis of Human Movement: Improving Performance in Sport and Exercise*. Human Kinetics.
- Koeneman, Margot a, Marieke W. Verheijden, Mai J. M. Chinapaw, and Marijke Hopman-Rock. 2011. "Determinants of Physical Activity and Exercise in Healthy Older Adults: A Systematic Review." *The international journal of behavioral nutrition and physical activity* 8:142.
- Kuhn, Max and Johnson, Kjell. 2013. "Data Pre-Processing." Pp. 27–59 in *Applied Predictive Modeling*. New York: Springer.
- Kwapisz, JR, GM Weiss, and SA Moore. 2011. "Activity Recognition Using Cell Phone Accelerometers." *ACM SigKDD Explorations Newsletter* 12(2):74–82.
- Lai, Daniel Tze Huei, Marimuthu Palaniswami, and Rezaul Begg. 2011. *Healthcare Sensor Networks: Challenges Toward Practical Implementation*. CRC Press.
- Lara, Oscar D., and Miguel a. Labrador. 2013. "A Survey on Human Activity Recognition Using Wearable Sensors." *IEEE Communications Surveys & Tutorials* 15(3):1192–1209.
- Ledger, Dan. 2014. *Inside Wearables: How the Science of Human Behavior Change*.
- Lee, MS, JG Lim, KR Park, and DS Kwon. 2009. "Unsupervised Clustering for Abnormality Detection Based on the Tri-Axial Accelerometer." Pp. 2–5 in *ICCA-SICE, 2009*.
- Lee, Myong-Woo, Adil Mehmood Khan, Ji-Hwan Kim, Young-Sun Cho, and Tae-Seong Kim. 2010. "A Single Tri-Axial Accelerometer-Based Real-Time Personal Life Log System Capable of Activity Classification and Exercise Information Generation." Pp. 1390–93 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Leon, Enrique, Graham Clarke, Victor Callaghan, and Faiyaz Doctor. 2010. "Affect-Aware Behaviour Modelling and Control inside an Intelligent Environment." *Pervasive and Mobile Computing* 6(5):559–74.
- Li, S. X. et al. 2010. "Validation of a New REM Sleep Behavior Disorder Questionnaire (RBDQ-HK)." *Sleep medicine* 11(1):43–48.
- Lin, Che-Wei, Ya-Ting C. Yang, Jeen-Shing Wang, and Yi-Ching Yang. 2012. "A Wearable Sensor Module with a Neural-Network-Based Activity Classification Algorithm for Daily Energy Expenditure Estimation." *IEEE transactions on*

- information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society* 16(5):991–98.
- Long, Xi, Bin Yin, and Ronald M. Aarts. 2009. "Single-Accelerometer-Based Daily Physical Activity Classification." Pp. 6107–10 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Lotfi, Ahmad, Caroline Langensiepen, Sawsan M. Mahmoud, and M. J. Akhlaghinia. 2011. "Smart Homes for the Elderly Dementia Sufferers: Identification and Prediction of Abnormal Behaviour." *Journal of Ambient Intelligence and Humanized Computing* 3(3):205–18.
- Lu, Ching-hu, and LC Fu. 2009. "Robust Location-Aware Activity Recognition Using Wireless Sensor Network in an Attentive Home." *IEEE Transactions on Automation Science and Engineering* 6(4):598–609.
- Ludwig, Wolfram et al. 2012. "Health-Enabling Technologies for the Elderly--an Overview of Services Based on a Literature Review." *Computer methods and programs in biomedicine* 106(2):70–78.
- Lühr, Sebastian, Geoff West, and Svetha Venkatesh. 2007. "Recognition of Emergent Human Behaviour in a Smart Home: A Data Mining Approach." *Pervasive and Mobile Computing* 3(2):95–116.
- De Maesschalck, R., D. Jouan-Rimbaud, and D. L. Massart. 2000. "The Mahalanobis Distance." *Chemometrics and Intelligent Laboratory Systems* 50(1):1–18.
- Mahmoud, Sawsan. 2013. "Behavioural Pattern Identification and Prediction in Intelligent Environments." *Applied Soft Computing* 13(4):1813–22.
- Mahoney, F. I., and D. W. Barthel. 1965. "Functional Evaluation: The Barthel Index." *Maryland state medical journal* 14:61–65.
- Marco, A. et al. 2008. "Location-Based Services for Elderly and Disabled People." *Computer Communications* 31(6):1055–66.
- Masse, Louise C. et al. 2005. "Accelerometer Data Reduction: A Comparison of Four Reduction Algorithms on Select Outcome Variables." *Medicine & Science in Sports & Exercise* 37(Supplement):S544–54.
- Mathie, Merryn J., Adelle C. F. Coster, Nigel H. Lovell, and Branko G. Celler. 2004. "Accelerometry: Providing an Integrated, Practical Method for Long-Term, Ambulatory Monitoring of Human Movement." *Physiological Measurement* 25(2):R1–20.

- Medjahed, Hamid, and Dan Istrate. 2009. "Human Activities of Daily Living Recognition Using Fuzzy Logic for Elderly Home Monitoring." Pp. 2001–6 in *IEEE International Conference on Fuzzy Systems*.
- Mintel. 2010. "Watches and Jewellery Retailing." Retrieved April 18, 2014 (<http://oxygen.mintel.com/display/479809/>).
- Molleda, Julio et al. 2012. "A Fast and Robust Decision Support System for in-Line Quality Assessment of Resistance Seam Welds in the Steelmaking Industry." *Computers in Industry* 63(3):222–30.
- Mubashir, Muhammad, Ling Shao, and Luke Seed. 2013. "A Survey on Fall Detection: Principles and Approaches." *Neurocomputing* 100:144–52.
- Murphy, Susan L. 2009. "Review of Physical Activity Measurement Using Accelerometers in Older Adults: Considerations for Research Design and Conduct." *Preventive medicine* 48(2):108–14.
- Navarro, Karla Felix, Elaine Lawrence, and Brian Lim. 2009. "Medical MoteCare: A Distributed Personal Healthcare Monitoring System." Pp. 25–30 in *2009 International Conference on eHealth, Telemedicine, and Social Medicine*. Ieee.
- Olmos Castillo, M^a Ángeles. 2008. "Tutorial de Introducción de Lógica Borrosa." Retrieved January 15, 2015 (<http://www.dma.fi.upm.es/java/fuzzy/tutfuzzy/indice.html>).
- Oppenheim, Alan V, Alan S. Willsky, and S. Hamid Nawab. 1983. *Signals and Systems*. edited by Englewood Cliffs. Prentice-Hall.
- Organización para la Cooperación y Desarrollo Económicos. 2012. "Gasto Medio En Salud En Los Países Pertenecientes a La OCDE." Retrieved January 9, 2015 (<http://www.oecd.org/centrodemexico/estadisticas/>).
- Orwat, Carsten, Andreas Graefe, and Timm Faulwasser. 2008. "Towards Pervasive Computing in Health Care - a Literature Review." *BMC medical informatics and decision making* 8(1):26.
- Oxley, Howard. 2009. "Policies for Healthy Ageing: An Overview." *OECD Health Working Papers* (42).
- Paladini, Luciana, Rick Hodder, Isabella Cecchini, Vincenzo Bellia, and Raffaele Antonelli Incalzi. 2010. "The MRC Dyspnoea Scale by Telephone Interview to Monitor Health Status in Elderly COPD Patients." *Respiratory medicine* 104(7):1027–34.

- Pantelopoulos, A., and N. G. Bourbakis. 2010. "A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 40(1):1–12.
- Papadopoulos, A., C. Crump, and B. Wilson. 2010. "Comprehensive Home Monitoring System for the Elderly." Pp. 214–15 in *Wireless Health 2010*.
- Parkka, J., and M. Ermes. 2006. "Activity Classification Using Realistic Data from Wearable Sensors." *IEEE Transactions on Information Technology in Biomedicine* 10(1):119–28.
- Patel, SA, and RP Benzo. 2007. "Activity Monitoring and Energy Expenditure in COPD Patients: A Validation Study." *COPD: Journal of Chronic Obstructive Pulmonary Disease* 4(2):107–12.
- Peng, Bo, Lei Zhang, and David Zhang. 2013. "A Survey of Graph Theoretical Approaches to Image Segmentation." *Pattern Recognition* 46(3):1020–38.
- Pérez, Antonio Solanas. 2005. *Estadística Descriptiva En Ciencias Del Comportamiento*. Thomson.
- Plasqui, G., a G. Bonomi, and K. R. Westerterp. 2013. "Daily Physical Activity Assessment with Accelerometers: New Insights and Validation Studies." *Obesity reviews : an official journal of the International Association for the Study of Obesity* 14(6):451–62.
- Popoola, O. P. 2012. "Video-Based Abnormal Human Behavior Recognition—A Review." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 42(6):865–78.
- Pyle, Dorian. 1999. *Data Preparation for Data Mining*. Morgan Kaufmann.
- Qian, Huimin, Yaobin Mao, Wenbo Xiang, and Zhiquan Wang. 2010. "Recognition of Human Activities Using SVM Multi-Class Classifier." *Pattern Recognition Letters* 31(2):100–111.
- Raad, M. W., and L. T. Yang. 2008. "A Ubiquitous Smart Home for Elderly." *Information Systems Frontiers* 11(5):529–36.
- Radhakrishna, R., Daniel Tobin, Mark Brennan, and Joan Thomson. 2012. "Ensuring Data Quality in Extension Research and Evaluation Studies." *Journal of Extension* 50(3):25–28.
- Rechel, Bernd et al. 2013. "Ageing in the European Union." *Lancet* 381(9874):1312–22.

- Reilly, John J. et al. 2003. "An Objective Method for Measurement of Sedentary Behavior in 3- to 4-Year Olds." *Obesity research* 11(10):1155–58.
- Scarmeas, Nikolaos et al. 2009. "Physical Activity, Diet, and Risk of Alzheimer Disease." *JAMA : the journal of the American Medical Association* 302(6):627–37.
- Schaar, AK, and Martina Ziefle. 2011. "Smart Clothing: Perceived Benefits vs. Perceived Fears." Pp. 601–8 in *2011 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*.
- Schoeller, D. a. 1988. "Measurement of Energy Expenditure in Free-Living Humans by Using Doubly Labeled Water." *The Journal of nutrition* 118(11):1278–89.
- Schreier, Günter, and Effie Amanatidou. 2014. "Personal Health Systems Research-Charting the European Landscape." *Studies in health technology and informatics* 200:179–81.
- Senthilkumaran, N., and R. Rajesh. 2009. "Edge Detection Techniques for Image Segmentation—a Survey of Soft Computing Approaches." *International Journal of Recent Trends in Engineering and Technology* 1(2):250–54.
- Shah, Manish N., Ryan McDermott, Suzanne M. Gillespie, Erin B. Philbrick, and Dallas Nelson. 2013. "Potential of Telemedicine to Provide Acute Medical Care for Adults in Senior Living Communities." *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine* 20(2):162–68.
- Sharma, Nandita, and Tom Gedeon. 2012. "Objective Measures, Sensors and Computational Techniques for Stress Recognition and Classification: A Survey." *Computer methods and programs in biomedicine* 108(3):1287–1301.
- Sharma, Neeraj, and Lalit M. Aggarwal. 2010. "Automated Medical Image Segmentation Techniques." *Journal of medical physics / Association of Medical Physicists of India* 35(1):3–14.
- Si, Xiao-sheng, Chang-hua Hu, Jian-bo Yang, and Zhi-jie Zhou. 2011. "A New Prediction Model Based on Belief Rule Base for System's Behavior Prediction." *IEEE Transactions on Fuzzy Systems* 19(4):636–51.
- SILHOUETTE Partnership. 2013a. *Join Potential for ICT-Based Innovations to Support Activity of the Elderly: Recommendations for Actions in the CREATOR Regions*.
- SILHOUETTE Partnership. 2013b. "SILHOUETTE: Using Novel Information and Communication Technologies for the Support of Elderly's Active Participation in

- the Information Society." Retrieved February 10, 2015 (<http://www.silhouette-project.eu/>).
- Singh, Aditya, Victor Lubecke, and Olga Boric-Lubecke. 2011. "Pulse Pressure Monitoring through Non-Contact Cardiac Motion Detection Using 2.45 GHz Microwave Doppler Radar." Pp. 4336–39 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Staudenmayer, John, David Pober, Scott Crouter, David Bassett, and Patty Freedson. 2009. "An Artificial Neural Network to Estimate Physical Activity Energy Expenditure and Identify Physical Activity Type from an Accelerometer." *Journal of applied physiology* 107(4):1300–1307.
- Steele, Robert, Amanda Lo, Chris Secombe, and Yuk Kuen Wong. 2009. "Elderly Persons' Perception and Acceptance of Using Wireless Sensor Networks to Assist Healthcare." *International journal of medical informatics* 78(12):788–801.
- Steele, Robert, K. Min, and Amanda Lo. 2012. "Personal Health Record Architectures: Technology Infrastructure Implications and Dependencies." *Journal of the American Society for Information Science and Technology* 63(6):1079–91.
- Taraldsen, Kristin et al. 2011. "Evaluation of a Body-Worn Sensor System to Measure Physical Activity in Older People with Impaired Function." *Physical therapy* 91(2):277–85.
- Taraldsen, Kristin, Sebastien F. M. Chastin, Ingrid I. Riphagen, Beatrix Vereijken, and Jorunn L. Helbostad. 2012. "Physical Activity Monitoring by Use of Accelerometer-Based Body-Worn Sensors in Older Adults: A Systematic Literature Review of Current Knowledge and Applications." *Maturitas* 71(1):13–19.
- Texas Instruments. 2014. "TI EZ430-Chronos." Retrieved April 21, 2014 (<http://processors.wiki.ti.com/index.php/EZ430-Chronos>).
- Thobbi, Anand, Ye Gu, and Weihua Sheng. 2011. "Using Human Motion Estimation for Human-Robot Cooperative Manipulation." Pp. 2873–78 in *IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Thomas, O. et al. 2010. "Wearable Sensor Activity Analysis Using Semi-Markov Models with a Grammar." *Pervasive and Mobile Computing* 6(3):342–50.
- Troosters, T., F. Sciruba, and S. Battaglia. 2010. "Physical Inactivity in Patients with COPD, a Controlled Multi-Center Pilot-Study." *Respiratory medicine* 104(7):1005–11.

- Trost, SG, and WK Wong. 2012. "Artificial Neural Networks to Predict Activity Type and Energy Expenditure in Youth." *Medicine and science in sports and exercise* 44(9):1801–9.
- Trost, Stewart G., Kerry L. Mciver, and Russell R. Pate. 2005. "Conducting Accelerometer-Based Activity Assessments in Field-Based Research." *Medicine and Science in Sports and Exercise* 37(Supplement):S531–43.
- United Nations, Department of Economic and Social Affairs, Population Division. 2013a. "Profiles of Ageing 2013." Retrieved January 9, 2015 (<http://esa.un.org/unpd/popdev/AgingProfiles2013/default.aspx>).
- United Nations, Department of Economic and Social Affairs, Population Division. 2013b. *World Population Ageing*.
- Vandewiele, Franck, and Cina Motamed. 2011. "An Unsupervised Learning Method for Human Activity Recognition Based on a Temporal Qualitative Model." P. 9 in *International Workshop on Behaviour Analysis and Video Understanding (ICVS 2011)*.
- Vanhees, Luc et al. 2005. "How to Assess Physical Activity? How to Assess Physical Fitness?" *European journal of cardiovascular prevention and rehabilitation* 12(2):102–14.
- Virone, Gilles. 2009. "Assessing Everyday Life Behavioral Rhythms for the Older Generation." *Pervasive and Mobile Computing* 5(5):606–22.
- Vogel, T. et al. 2009. "Health Benefits of Physical Activity in Older Patients: A Review." *International journal of clinical practice* 63(2):303–20.
- De Vries, Nienke M., J. Bart Staal, Marcel G. M. Olde Rikkert, and Maria W. G. Nijhuis-van der Sanden. 2013. "Evaluative Frailty Index for Physical Activity (EFIP): A Reliable and Valid Instrument to Measure Changes in Level of Frailty." *Physical therapy* 93(4):551–61.
- VTI Technologies. 2009. "Hoja de Características Del Acelerómetro CMA3000-D01." Retrieved April 20, 2014 (<http://www.alldatasheet.es/datasheet-pdf/pdf/313334/VTI/CMA3000-D01.html>).
- Wang, Liang, Tao Gu, Xianping Tao, Hanhua Chen, and Jian Lu. 2011. "Recognizing Multi-User Activities Using Wearable Sensors in a Smart Home." *Pervasive and Mobile Computing* 7(3):287–98.
- Ward, Dianne S., Kelly R. Evenson, Amber Vaughn, Anne Brown Rodgers, and Richard P. Troiano. 2005. "Accelerometer Use in Physical Activity: Best Practices

- and Research Recommendations." *Medicine & Science in Sports & Exercise* 37(Supplement):S582–88.
- Ward, Jamie A., Paul Lukowicz, and Hans W. Gellersen. 2011. "Performance Metrics for Activity Recognition." *ACM Transactions on Intelligent Systems and Technology* 2(1):1–23.
- Warren, Janet M. et al. 2010. "Assessment of Physical Activity - a Review of Methodologies with Reference to Epidemiological Research: A Report of the Exercise Physiology Section of the European Association of Cardiovascular Prevention and Rehabilitation." *European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology* 17(2):127–39.
- Waschki, Benjamin et al. 2012. "Physical Activity Monitoring in COPD: Compliance and Associations with Clinical Characteristics in a Multicenter Study." *Respiratory medicine* 106(4):522–30.
- Weiss, Allison R., Nathan L. Johnson, Nathan a Berger, and Susan Redline. 2010. "Validity of Activity-Based Devices to Estimate Sleep." *Journal of clinical sleep medicine* 6(4):336–42.
- Welk, Gregory J. 2002. *Physical Activity Assessments for Health-Related Research*.
- Welk, Gregory J., James McClain, and Barbara E. Ainsworth. 2012. "Protocols for Evaluating Equivalency of Accelerometry-Based Activity Monitors." *Medicine and science in sports and exercise* 44(1 Suppl):S39–49.
- Wild, K., L. Boise, J. Lundell, and A. Foucek. 2008. "Unobtrusive in-Home Monitoring of Cognitive and Physical Health: Reactions and Perceptions of Older Adults." *Journal of Applied Gerontology* 27(2):181–200.
- Wu, Pang, HK Peng, Jiang Zhu, and Ying Zhang. 2012. "Senscare: Semi-Automatic Activity Summarization System for Elderly Care." Pp. 1–19 in *Mobile Computing, Applications, and Services*.
- Xiang, Tao, and Shaogang Gong. 2008. "Incremental and Adaptive Abnormal Behaviour Detection." *Computer Vision and Image Understanding* 111(1):59–73.
- Xie, Miao, Song Han, Biming Tian, and Sazia Parvin. 2011. "Anomaly Detection in Wireless Sensor Networks: A Survey." *Journal of Network and Computer Applications* 34(4):1302–25.

- Yamana, Yusuke et al. 2011. "A Sensor for Monitoring Pulse Rate, Respiration Rhythm, and Body Movement in Bed." Pp. 5323–26 in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- Yang, Jun. 2009. "Toward Physical Activity Diary: Motion Recognition Using Simple Acceleration Features with Mobile Phones." Pp. 1–9 in *1st international workshop on Interactive multimedia for consumer electronics*.
- Ye, Juan, Simon Dobson, and Susan McKeever. 2012. "Situation Identification Techniques in Pervasive Computing: A Review." *Pervasive and Mobile Computing* 8(1):36–66.
- Yuan, Qilong, and I-ming Chen. 2012. "Human Velocity and Dynamic Behavior Tracking Method for Inertial Capture System." *Sensors and Actuators A: Physical* 183:123–31.
- Zatsiorsky, Vladimir M. 2002. *Kinetics of Human Motion*. Human Kinetics.
- Zhang, Shaoyan, Alex V Rowlands, Peter Murray, and Tina L. Hurst. 2012. "Physical Activity Classification Using the GENE Wrist-Worn Accelerometer." *Medicine and science in sports and exercise* 44(4):742–48.
- Zhou, Zhongna, Xi Chen, and YC Chung. 2008. "Activity Analysis, Summarization, and Visualization for Indoor Human Activity Monitoring." *IEEE Transactions on Circuits and Systems for Video Technology* 18(11):1489–98.
- Zhu, Chun, and Weihua Sheng. 2010. "Recognizing Human Daily Activity Using a Single Inertial Sensor." Pp. 282–87 in *2010 8th World Congress on Intelligent Control and Automation*. Ieee.
- Zhu, Chun, and Weihua Sheng. 2011. "Motion and Location-Based Online Human Daily Activity Recognition." *Pervasive and Mobile Computing* 7(2):256–69.
- Ziefle, Martina, and Carsten Rocker. 2010. "Acceptance of Pervasive Healthcare Systems: A Comparison of Different Implementation Concepts." Pp. 1–6 in *4th International ICST Conference on Pervasive Computing Technologies for Healthcare*. Ieee.

APÉNDICE A.

ABREVIATURAS, GLOSARIO Y NOMENCLATURA

A.1. ABREVIATURAS Y SIGLAS

<i>ASN</i>	Ambient sensor network
<i>BSN</i>	Body sensor network
<i>DAFO</i>	Debilidades, amenazas y fortalezas
<i>GDP</i>	Gross domestic product; producto interior bruto (PIB)
<i>HDMI</i>	High-definition multimedia interface
<i>HL7</i>	Health level seven
<i>ADL</i>	Activity daily living; actividades de la vida diaria
<i>IIR</i>	Infinite impulse response
<i>JIM</i>	Jerk-based information magnitude
<i>PC</i>	Personal computer
<i>PIR</i>	Passive infrared sensor
<i>PPP</i>	Purchasing power parity; paridad del poder adquisitivo (PPA)
<i>RFID</i>	Radio frequency identification
<i>REST</i>	Representational state transfer
<i>RF</i>	Radio frecuencia
<i>SD</i>	Standard deviation

<i>SMA</i>	Signal magnitude area
<i>SoC</i>	System on a chip
<i>SQL</i>	Structure query language
<i>TIC</i>	Tecnologías de la información y la comunicación

A.2. GLOSARIO DE TÉRMINOS

<i>Actividad física</i>	Se define formalmente como cualquier movimiento del cuerpo humano que se traduce en un gasto energético
<i>Adaptativo</i>	Hace referencia a la capacidad de adaptación del método propuesto. Un método adaptativo es aquel capaz de modificar su funcionamiento en función de determinadas características. En este caso, el método es adaptativo al usuario y al instante temporal en el que se trabaja
<i>Arquitectura de monitorización</i>	Elementos hardware y software necesarios para llevar a cabo la monitorización de usuarios. Se identifican tres elementos: elemento de medida, elemento concentrador y elemento remoto
<i>Automático</i>	Hace referencia a la capacidad de un método de funcionar por sí solo sin necesidad de configuración humana
<i>Autónomo</i>	Hace referencia a la capacidad de un método de funcionar sin errores y sin necesidad de interacción humana durante un periodo de tiempo prolongado
<i>Comportamiento anómalo</i>	Se produce cuando durante un tramo temporal se registran niveles de actividad no esperados de acuerdo al comportamiento esperado o rutina
<i>Computación vestible</i>	Traducción del término inglés <i>wearable computing</i> . También nombrado en algunos trabajos como <i>computación portable</i> . Hace referencia a la integración de elementos de computación (sensores, actuadores, procesadores...) en prendas de vestir o complementos vestibles
<i>Cuentas</i>	Traducción del término inglés <i>counts</i> . Hace referencia a los valores obtenidos directamente de la cuantificación analógica-digital que tiene lugar en un acelerómetro. Su uso simplifica los cálculos en el procesamiento de las medidas

<i>Cuidador informal</i>	Se denominan cuidadores informales al conjunto de personas sin especialización sanitaria que dedican una importante actividad diaria al cuidado de personas dependientes (generalmente familiares)
<i>Fiabilidad</i>	La fiabilidad en la detección se mide en este trabajo a partir del cálculo de la precisión y la sensibilidad del método de detección propuesto. Un método fiable en la detección será aquel que presenta una alta precisión junto con una alta sensibilidad
<i>Método supervisado</i>	Se trata de un método que requiere una etapa previa de entrenamiento para su funcionamiento. Esta etapa previa de entrenamiento suele estar relacionada con una etapa de etiquetado de datos a partir de la cual se realiza el ajuste y diseño final del sistema
<i>No Invasivo</i>	Hace referencia a la capacidad de un elemento de pasar inadvertido y de funcionar de forma transparente para el usuario y su entorno
<i>Precisión</i>	En un método de detección, la precisión se calcula mediante la fracción de eventos detectados correctamente respecto al total de eventos detectados por el método. Representa el porcentaje de la salida del método que es correcta
<i>Requisito funcional</i>	Se trata de un requisito o requerimiento definido a partir de una funcionalidad que un sistema debe cumplir
<i>Requisito no funcional</i>	Se trata de un requisito o requerimiento que un sistema debe cumplir y que no hacen referencia a su funcionalidad. Dentro de los requisitos no funcionales destacan los requisitos técnicos y los requisitos de usuario
<i>Rutina</i>	Es el comportamiento esperado según las costumbres y hábitos de un usuario
<i>Sensibilidad</i>	También conocida como <i>recall</i> , exhaustividad o cobertura. En un método de detección, la sensibilidad representa la fracción de eventos detectados respecto al total de eventos ocurridos. Representa el porcentaje de eventos que son detectados por el método

A.3. NOMENCLATURA

<i>a</i>	Medida de la aceleración. Se referencia como $a_{gravity}$ a su componente gravitacional y a_{body} a la componente relacionada con el movimiento del cuerpo humano. Se utilizan los subíndices x , y , z para representar las componentes de cada eje del movimiento
<i>Actividad evaluable</i>	Se refiere a la variable <i>Actividad física acondicionada</i> en el día sobre el que se desea llevar a cabo la detección de anomalías de comportamiento
<i>Actividad física</i>	Hace referencia a la actividad física medida mediante el estimador JIM durante un día completo de monitorización. Puesto que las medidas se toman cada minuto, la variable tendrá 1440 muestras
<i>Actividad física acondicionada</i>	Salida del método de acondicionamiento propuesto (primera etapa del método de modelado). Se calcula sobre la variable <i>Actividad física</i> y su expresión se describe en la Ecuación 5.3
<i>Comportamiento anómalo</i>	Variable de salida del método de detección (salida de la fase III). Es la transformación de <i>Valor de anomalía</i> tras pasar por un proceso de filtrado de parámetros N_L y V
<i>Diferencia de actividad</i>	Cálculo de la diferencia entre la <i>Rutina</i> y la <i>Actividad física</i> en el día sobre el que se aplica el método de detección. Su expresión se muestra en la Ecuación 6.2 y puede interpretarse como el cálculo de los residuos de un modelo
<i>Diferencia de actividad filtrada</i>	Es la salida de la fase I del método de detección. Su cálculo se representa en la Ecuación 6.3
<i>Jerk</i>	Derivada de la aceleración. Se utilizan los subíndices x , y , z para representar el cálculo en cada uno de los ejes del movimiento
<i>JIM</i>	Medida de la actividad física utilizando el estimador JIM propuesto en este trabajo. Su expresión matemática se presenta en la Ecuación 4.3
<i>N</i>	Utilizado en varias expresiones para representar el número total de muestras sobre el que se realiza un cálculo

N_L	Número de muestras utilizado en el proceso de filtrado de anomalías de la fase III del método de detección. Este parámetro fija el tamaño para la detección de comportamientos anómalos en la variable <i>Valor de anomalía</i> . Se propone el uso de $N_L = 60$ al trabajar con resolución horaria
N_m	Número de muestras relacionado con la resolución que se desea en el método de detección. Se propone el valor 60 ($N_m = 60$) para tener una resolución a nivel de hora en la detección
N_{min}	Mínima cantidad de datos utilizada en el método de acondicionamiento (ver sección 5.3.2.2) para determinar si debe utilizarse una estrategia de eliminación o de auto-completado. Se propone la eliminación de un día si hay menos de 17 horas de monitorización ($N_{min} = 1020$)
N_r	Número de días utilizado para el cálculo de <i>Rutina</i> . Se propone el uso de 5 días ($N_r = 5$)
<i>Rutina</i>	Variable que representa la rutina identificada mediante el método de modelado propuesto. Su expresión matemática se presenta en la Ecuación 5.4 y para un día d se calcula a partir de la <i>Actividad física</i> medida en los N_r días anteriores (desde el día $d-N_r$ hasta el día $d-1$)
SD	Representa a la desviación estándar de una variable. Con el subíndice rutina representa la desviación sobre la variable <i>Rutina</i> (ver Ecuación 6.3) y con el subíndice evaluable la desviación sobre la <i>Actividad evaluable</i> (ver Ecuación 6.4). El cálculo se realiza mediante una ventana móvil de tamaño N_m y es la salida de la fase II del método de detección propuesto
SMA	Medida de la actividad física utilizando el estimador más utilizado en el estado del arte. Su expresión matemática se muestra en la Ecuación 4.1
T	Periodo de tiempo durante el cual se realiza el muestreo del acelerómetro para la estimación de la actividad física

V	Puntuación mínima de anomalía para que las muestras de la variable <i>Valor de anomalía</i> sean consideradas como parte de un comportamiento anómalo. Este parámetro se utiliza en el proceso de filtrado de anomalías de la fase III del método de detección. Con $V = 0$ se detectan todos los cambios de actividad y con $V = 1$ solo los comportamientos anómalos
<i>Valor de anomalía</i>	Puntuación dada por la función Fuzzy de evaluación de anomalías. Es un valor entre -1 y 1 que indica el grado y sentido de la anomalía. Valores en el intervalo $[-1, 0)$ representan decrementos anómalos de actividad, valores en el intervalo $(0, 1]$ representan incrementos anómalos de actividad y un valor 0 representa la normalidad. El cálculo se representa en la Ecuación 6.6 para cada muestra de actividad. La variable <i>Valor de anomalía</i> tendrá 1440 muestras, igual que la <i>Actividad física</i>

APÉNDICE B.

RECOLECCIÓN DE REQUISITOS DE USUARIO

Este apéndice recoge el cuestionario utilizado durante el proyecto SILHOUETTE (SILHOUETTE Partnership 2013b) para la identificación de los principales problemas que la adopción de las TIC plantea en las personas mayores. Los resultados obtenidos mediante el presente cuestionario se describen en la sección 2.5.2 del Capítulo 2.

Proyecto SILHOUETTE – Cuestionario PARTE 1

1. Género (por favor, marque la casilla que corresponda)

- Hombre
 Mujer

2. ¿En qué año nació Ud?

3. ¿Cuál es su nivel de estudios? (por favor, marque la casilla que corresponda)

- Escuela
 Estudios básicos/primarios (hasta los 14 años)
 Bachiller elemental
 Bachiller superior
 Universidad
 Oficialía
 Maestría / Capataz
 Otros (por favor, especifique):

4. Desde el primer año de educación primaria, ¿durante cuántos años estuvo usted estudiando aproximadamente?

 Años

5. ¿Dónde se encuentra su domicilio habitual?

- En una ciudad
 En un pueblo
 En una casa aislada, alejada de otras casas

6. ¿Vive usted con alguien? (marido, esposa, compañero, hijos,...)

- Sí
 No

7. ¿Cuántas personas viven en la misma casa que usted de forma habitual? (usted incluido)

 personas

8. ¿Vive usted en un piso o en una casa?

- En un piso
 En una casa
 Otro (por favor, especifique cuál):

9. ¿Es usted propietario de la casa o piso en el que vive?

- Sí, soy propietario
 No, estoy de alquiler
 No soy propietario pero tampoco pago ningún alquiler (a efectos prácticos, alojamiento gratuito)

10. En caso de que viva de alquiler, el propietario del lugar en el que usted vive es...:

- Un propietario privado
 Una entidad pública

11. ¿Le gustaría cambiar su lugar de residencia o las condiciones del lugar en el que reside?

- No me gustaría cambiar mi lugar de residencia (hogar) y tampoco veo necesaria ninguna remodelación o renovación en mi casa actual, está bien tal cual está y no me plantea ningún problema en relación a mis necesidades. Me gustaría seguir viviendo en mi casa, pero renovando o remodelando algunas cosas que me ayuden y me faciliten las labores del día a día.
 Me gustaría cambiar mi lugar de residencia a otro que se adecúe mejor a mis necesidades actuales.

12. En caso de necesidad derivada de una pérdida de autonomía, ¿ha considerado alguna alternativa a su lugar y condiciones actuales?

- Sí
 No

13. En caso afirmativo, ¿qué tipo de solución consideraría o ha considerado? (puede seleccionar más de una casilla si lo desea)

- Alojarme dentro del entorno familiar (p.e. vivir en casa de algún familiar).
 Irme a vivir a una residencia o a algún tipo de alojamiento vigilado.
 Seguir viviendo en mi casa pero contando con servicios de atención en el hogar.
 Seguir viviendo en mi casa pero con asistencia basada en nuevas tecnologías (p.e. telecuidados, detectores de caídas, etc).

Apéndice B

14. ¿Tiene usted alguna ocupación? (puede marcar más de una casilla si lo desea)

<input type="checkbox"/>	Sí, actualmente estoy en activo (trabajo remunerado por cuenta ajena)
<input type="checkbox"/>	Sí, actualmente estoy en activo (trabajo remunerado por cuenta propia)
<input type="checkbox"/>	Trabajo de forma ocasional
<input type="checkbox"/>	No, estoy desempleado
<input type="checkbox"/>	No, estoy (pre)jubilado
<input type="checkbox"/>	Otro (por favor, especificar): _____

15. ¿Cuál es su nivel de ingresos mensuales?

<input type="checkbox"/>	Menos de 600 €
<input type="checkbox"/>	Entre 600 y 1 000 €
<input type="checkbox"/>	Entre 1 000 y 1 500 €
<input type="checkbox"/>	Entre 1 500 y 2 000 €
<input type="checkbox"/>	Más de 2 000 €

16. En la última semana, ¿ha estado (se ha visto usted) con...

	Sí	No
... familiares o parientes?	<input type="checkbox"/>	<input type="checkbox"/>
... amigos o conocidos?	<input type="checkbox"/>	<input type="checkbox"/>

17. En los últimos 6 meses, ¿cuántas veces ha ido al médico aproximadamente?

Veces

18. ¿Alguna vez ha hecho algún curso relacionado con informática u ordenadores?

Sí
 No

19. ¿Está usted de acuerdo con las siguientes afirmaciones?

Por favor, marque con un círculo la respuesta que más se adecúe a su opinión.

	Totalmente en desacuerdo	Mayormente en desacuerdo, pero no del todo	Ni de acuerdo ni en desacuerdo	De acuerdo en algunas cosas pero no en todas	Totalmente de acuerdo
Me siento solo frecuentemente.	1	2	3	4	5
Necesito ayuda para llevar a cabo las actividades del día a día.	1	2	3	4	5
Creo que tendré que cambiar de casa en los próximos años por cuestiones de salud.	1	2	3	4	5
Me resulta difícil tener que hacer la compra solo.	1	2	3	4	5
Los médicos están disponibles cuando necesito alguna consulta.	1	2	3	4	5
No hablo con nadie.	1	2	3	4	5
Habitualmente olvido cosas.	1	2	3	4	5
No tengo ayuda suficiente para hacer las tareas del hogar.	1	2	3	4	5
Estoy satisfecho con mis condiciones de vida actuales.	1	2	3	4	5
Mi hogar es adecuado teniendo en cuenta mis condiciones de salud.	1	2	3	4	5
Cuando estoy solo, me alimento de forma adecuada.	1	2	3	4	5

3

4

20. ¿Está usted de acuerdo con las siguientes afirmaciones?

Por favor, marque con un círculo la respuesta que más se adecúe a su opinión.

	Totalmente en desacuerdo	Mayormente en desacuerdo, pero no del todo	Ni de acuerdo ni en desacuerdo	De acuerdo en algunas cosas pero no en todas	Totalmente de acuerdo
Cuando no uso internet me siento aislado de la gente joven.	1	2	3	4	5
Me da vergüenza no saber utilizar un ordenador ni Internet.	1	2	3	4	5
Las personas que son importantes para mí usan el ordenador e Internet.	1	2	3	4	5
Los dispositivos electrónicos son peligrosos para la salud.	1	2	3	4	5
Las nuevas tecnologías reducen la vida familiar.	1	2	3	4	5
Cuando uno utiliza Internet y los teléfonos móviles, no puede proteger su vida privada.	1	2	3	4	5
Estaría más interesado en las Tecnologías de la Información y la Comunicación si fuesen menos caras.	1	2	3	4	5
Tengo información suficiente sobre cómo funcionan los dispositivos electrónicos y los ordenadores.	1	2	3	4	5
No puedo usar dispositivos electrónicos porque los manuales son muy difíciles de entender.	1	2	3	4	5
Las Tecnologías de la Información y las Comunicaciones son necesarias cuando uno vive solo.	1	2	3	4	5
Estaría más interesado en usar dispositivos electrónicos y ordenadores si algún profesional me enseñase a manejarlos.	1	2	3	4	5
Hay tantos modelos distintos de cada tipo de dispositivo electrónico que no sabría cuál elegir.	1	2	3	4	5
Los manuales de los dispositivos usan demasiadas palabras técnicas.	1	2	3	4	5
Los precios de los dispositivos electrónicos más habituales están, en términos generales, al alcance de mi bolsillo.	1	2	3	4	5

5

6

(sigue...)	Totalmente en desacuerdo	Mayormente en desacuerdo, pero no del todo	Ni de acuerdo ni en desacuerdo	De acuerdo en algunas cosas pero no en todas	Totalmente de acuerdo
Sería útil que las personas retiradas tuvieran a su disposición una oficina de información y asistencia específicamente diseñada para guiarlos y acompañarlos en la compra y uso de equipos electrónicos y ordenadores.	1	2	3	4	5
Cuando un dispositivo electrónico u ordenador no funciona bien, ¿cree que los departamentos de post-venta y ayuda al cliente son satisfactorios?	1	2	3	4	5
No confío en los dispositivos electrónicos ni en los proveedores de servicios porque ellos juegan con ventaja debido a mi falta de conocimientos (pueden hacerme comprar dispositivos inútiles para mí).	1	2	3	4	5
Preferiría dispositivos electrónicos y/o ordenadores equipados con pantallas táctiles (dispositivos que pueden ser manejados pulsando directamente en la pantalla).	1	2	3	4	5
Las teclas de los teléfonos móviles son muy pequeñas para escribir y enviar mensajes.	1	2	3	4	5
Los dispositivos electrónicos con múltiples funciones (p.e. un teléfono móvil que también me sirve para oír la radio) son más prácticos que los dispositivos que sólo tienen una función.	1	2	3	4	5
Los dispositivos electrónicos usan términos técnicos en inglés que no entiendo y que influyen en que no los pueda utilizar.	1	2	3	4	5
Los manuales de usuario de dispositivos electrónicos y ordenadores están escritos en letra muy pequeña que no puedo leer bien, lo que me desanima a leerlos y por tanto a aprender a utilizar bien los dispositivos.	1	2	3	4	5

Proyecto SILHOUETTE – Cuestionario PARTE 2

1. ¿Tiene usted ordenador?

Sí
 No – Si ha marcado esta opción, vaya por favor hasta la pregunta número 7 y continúe el cuestionario desde ahí

2. Su ordenador, ¿es algún modelo simplificado diseñado específicamente para personas mayores?

Sí
 No

3. ¿Tiene usted en casa conexión a Internet?

Sí
 No – Si ha marcado esta opción, vaya por favor hasta la pregunta número 7 y continúe el cuestionario desde ahí

4. ¿Tiene usted alguna cuenta de correo electrónico?

Sí
 No

5. Durante el mes pasado, ¿se conectó usted a través de Internet con algún...

	Sí	No
... familiar o pariente?		
... amigos o conocidos?		
... desconocido hasta ese momento?		

6. En el último mes, ¿cuántos días utilizó usted Internet en su casa?

Días

7

7. Si usted tiene Internet, O, si usted tuviese la posibilidad de tener Internet en su casa, ¿la usaría o le gustaría usarla para...

	Sí	No	No lo sé
... conocer gente nueva?			
... tener contacto más frecuente con sus hijos, nietos o amigos?			
... consultar las "páginas amarillas" y otros anuarios?			
... encontrar información educativa o profesional relacionada con el mercado de trabajo?			
... encontrar información, leer sus periódicos y revistas favoritos?			
... monitorizar sus acciones en la bolsa de valores?			
... conectarse a Internet?			
... hacer compras en remoto?			
... organizar sus viajes y/o vacaciones?			
... llevar a cabo tareas administrativas (por ej. anotar una cita)?			
... gestionar sus cuentas bancarias?			
... pagar facturas?			
... encontrar recetas de cocina?			
... encontrar información médica y relacionada con la salud?			
... registrarse en alguna red social, como por ej. Facebook?			

8. ¿Tiene Ud. su propio teléfono móvil?

Sí
 No – Si ha marcado esta opción, vaya por favor hasta la pregunta número 11 y continúe el cuestionario desde ahí

9. Si lo tiene, ¿es algún modelo simplificado, diseñado específicamente para gente mayor?

Sí
 No – Si ha marcado esta opción, vaya por favor hasta la pregunta número 11 y continúe el cuestionario desde ahí

10. Si usted tiene su propio teléfono móvil simplificado, ¿tiene algún tipo de tarifa o condiciones especiales (plan de pago) asociadas a dicho tipo de móvil?

Sí
 No

8

11. ¿Tiene usted alguno de los siguientes dispositivos electrónicos?

Televisión
 Telecuidados
 Detector de caídas
 Teléfono fijo
 Impresora
 Cámara / videocámara
 Fax
 Videoconsola de juegos (en caso afirmativo, ¿cuál?)
 Radio
 Ninguno de los anteriores
 Otros (especificar):

12. En el próximo año, ¿se plantea reemplazar o comprar alguno de los siguientes dispositivos?

Televisión
 Telecuidados
 Teléfono móvil
 Detector de caídas
 Teléfono fijo
 Suscripción a Internet
 Ordenador
 Impresora
 Cámara / Videocámara
 Fax
 Videoconsola de juegos (en caso afirmativo, ¿cuál?)
 Radio
 Ninguno de los anteriores
 No lo sé
 Otros (especificar):

9

13. ¿Cuáles son las necesidades por las que un dispositivo electrónico sencillo de utilizar podría ser interesante para usted?

Llevar a cabo más actividades de tiempo libre
 Salud / telemedicina / diagnóstico de salud en remoto
 Seguridad
 Comunicación
 Ejercicio físico / deportes
 Ayuda en las tareas del hogar (por ej. una aspiradora)
 Moverme dentro de casa (por ej. salvaescaleras eléctrico para casa)
 Moverme fuera de casa
 Otras (especificar):

14. Aproximadamente, ¿cuánto gasta usted al mes en...

... teléfono móvil? €
 ... televisión (canales de pago, cable, satélite)? €
 ... teléfono fijo? €
 ... Internet? €
 ... servicio triple: "suscripción a Internet + teléfono fijo + televisión"? €

15. ¿Necesita usted ayuda para llevar a cabo las siguientes actividades? (si nunca lo ha intentado responda "No lo sé")

	Sí	No	No lo sé
Configurar los canales de televisión			
Teclar un texto en el ordenador			
Enviar un mensaje de texto por teléfono			
Enviar fotografías con su teléfono móvil			
Leer sus e-mails y enviar una respuesta			
Imprimir desde un ordenador			
Encontrar una página web concreta			
Descargarse de Internet documentos e imágenes			
Llamar a amigos con skype (o algún programa similar)			
Adjuntar documentos a e-mails que quiera enviar			

FIN DEL CUESTIONARIO. ¡ MUCHAS GRACIAS POR SU COLABORACION !

10

APÉNDICE C.

SISTEMAS DE MONITORIZACIÓN DE PERSONAS MAYORES

C.1. RESULTADOS DEL PROYECTO SILHOUETTE

Este apéndice recoge los principales resultados obtenidos en la ejecución del proyecto internacional SILHOUETTE (SILHOUETTE Partnership 2013b), cofinanciado por INTERREG IVC y FEDER a través del mini-programa CREATOR. En él se identificaron y compartieron prácticas, conocimientos y herramientas TIC en 7 regiones europeas con el objetivo de promover la participación de las personas mayores en la sociedad de la información.

La Tabla C.1 describe los principales resultados identificando el nombre e imagen del producto, el país en el que se comercializa y su coste aproximado

Tabla C.1. Resultados obtenidos en la ejecución del proyecto SILHOUETTE

Producto	País	Coste	Descripción	Imagen
Telecare	Francia	20-32€ mes	Servicio formado por un botón del pánico y un dispositivo transmisor para facilitar la comunicación entre los usuarios y un centro de atención	
Geo location	Francia	30-40€ mes	A través de un dispositivo en forma de reloj se puede controlar la localización del usuario. Se permite definir un perímetro de seguridad para que el sistema envíe una alarma a familiares/expertos cuando el usuario lo sobrepasa	
Fall detector	Francia	40-45€ mes	Detector de caídas que envía una señal de alarma cuando detecta una posible caída del usuario seguida de 15 segundos de inactividad. Se genera de forma automática una llamada	
Mini phone	Francia	25-38€ mes	Dispositivo en forma de brazalete que es en realidad un teléfono móvil disponible 24/7 y que permite programar hasta 11 números de teléfono a los que se llama de forma automática cuando sea preciso. El dispositivo es resistente al agua y está pensado para que el usuario lo porte en todo momento	

Producto	País	Coste	Descripción	Imagen
Safety one solution	Finlandia	15-30€ mes	Dispositivo similar al botón del pánico. Se comercializa como un brazalete con un botón integrado que el usuario puede pulsar cuando necesite asistencia	
Button of life	Polonia	20-25€ mes	Similar al botón del pánico. Los usuarios pueden pulsar un botón para avisar a los servicios médicos de una emergencia	
e-Troska	Polonia	175€ mes	Servicio proporcionado por la compañía e-Troska para las personas mayores de Polonia. Se trata de un concepto similar al botón del pánico	
Check Up Care	Suecia	-	Se trata de un dispositivo compuesto por sencillos sensores biomédicos que pueden ser utilizados por las personas mayores. Permite la monitorización remota de varios parámetros fisiológicos: nivel de azúcar en sangre, hemoglobina, presión sanguínea, pulso, peso... La información se envía a un servidor de forma inalámbrica permitiendo su acceso tanto por familiares como por profesionales del ámbito sanitario	

Producto	País	Coste	Descripción	Imagen
Safety alarm	Suecia	11€ mes	Puede tener tanto forma de brazalete como forma de collar. Integran un botón que al ser pulsado por el usuario se envía una alarma indicando que el usuario precisa de asistencia	
Botón de alarma	España	-	Dispositivo que se proporciona de forma gratuita bajo preinscripción médica con colaboración de Cruz Roja. El usuario genera una alarma requiriendo asistencia cuando pulsa un botón	

C.2. SOLUCIONES COMERCIALES

Este Apéndice recoge las principales soluciones comerciales existentes tanto a nivel europeo como a nivel mundial para la monitorización de personas. La Tabla C.2 muestra el nombre, fabricante, descripción e imagen de las principales soluciones existentes. En el apartado categoría, se indica el tipo de servicio de acuerdo a la clasificación mostrada en la Tabla 3.2.

Tabla C.2. Soluciones comerciales para la monitorización de personas

Producto	Descripción	Imagen
Health buddy¹	Cada día el dispositivo indica al usuario que medidas bio-médicas es necesario tomar. En la pantalla presenta consejos para mejorar su estilo de vida. El dispositivo recolecta los datos y los envía a un servidor central al cual expertos en medicina tienen acceso Categoría: <i>Evaluación del estado de salud</i>	

¹ BOSCH Healthcare: Health Buddy System. URL: http://www.bosch-telehealth.com/en/us/products/health_buddy/health_buddy.html. Acceso: 15/02/2014.

Producto	Descripción	Imagen
Tele station²	<p>Medición de variables médicas a través de sensores de salud. La información se envía a los expertos médicos a través de una línea de teléfono</p> <p>Categoría: <i>Evaluación del estado de salud</i></p>	
Hommed³	<p>Solución integral compuesta por varios dispositivos y servicios. Proporcionan, entre otros, dispositivos para la monitorización de pacientes a través del control de variables médicas, dispositivos para transmitir recordatorios e interfaz de comunicación (audio/texto) entre el paciente y el experto sanitario</p> <p>Categoría: <i>Evaluación del estado de salud, Motivación y recordatorios, Educación y consulta</i></p>	
Intel health guide⁴	<p>Combinación de un dispositivo a través del cual el usuario mide sus variables médicas con la posibilidad de acceder a contenido educacional para mejorar el manejo de enfermedades por parte del paciente</p> <p>Categoría: <i>Evaluación del estado de salud, Educación y consulta</i></p>	
Ideal LIFE Pod⁵	<p>Dispositivos conectados a un concentrador para la recolección y transmisión de variables médicas</p> <p>Categoría: <i>Evaluación del estado de salud</i></p>	

² PHILIPS TeleStation: Healthy at home. URL:

http://www.healthcare.philips.com/us_en/products/telehealth/products/telestation.wpd. Acceso: 15/02/2014.

³ Honeywell HomMed. URL: <http://www.hommed.com>. Acceso: 15/02/2014.

⁴ Intel Across Healthcare and Life Science Solutions. URL:

<http://www.intel.com/content/www/us/en/healthcare-it/healthcare-overview.html>. Acceso: 15/02/2014.

⁵ Ideal LIFE Pod. URL: <http://www.ideallifeonline.com/products/>. Acceso: 15/02/2014.

Producto	Descripción	Imagen
<p>Health any where⁶</p>	<p>Monitorización de variables médicas a través de sensores inalámbricos. El sistema se combina con una plataforma multidispositivo que permite la creación de recordatorios y alertas (tablet, smartphone) al que pueden acceder usuarios y expertos.</p> <p>Categorías: <i>Evaluación del estado de salud, Motivación y recordatorios, Educación y consulta</i></p>	
<p>Comfort zone⁷</p>	<p>Utilizando GPS y comunicaciones móviles se monitoriza la localización de la persona que porta el dispositivo. A través de una interfaz web los familiares pueden monitorizar la localización del usuario e indicar perímetros de seguridad recibiendo un aviso cuando el usuario los sobrepasa</p> <p>Categoría: <i>Detección de condiciones adversas</i></p>	
<p>Quiet care⁸</p>	<p>Solución domótica basada en la integración de sensores de presencia en el hogar del usuario monitorizado. El sistema detecta situaciones anómalas y envía avisos a los cuidadores para que intervengan si lo consideran oportuno</p> <p>Categoría: <i>Detección de condiciones adversas</i></p>	

⁶ AnyWhereHealthCare. URL: <http://www.anywherehealthcare.com.au>. Acceso: 15/02/2014.

⁷ Alzheimer's Association Comfort Zone Check-In. URL: <http://www.comfortzonecheckin.com/WhatIsComfortZoneCheckIn.aspx>. Acceso: 15/02/2014.

⁸ Care Innovations QuietCare. URL: <http://www.careinnovations.com/products/quietcare-assisted-living-technology>. Acceso: 15/02/2014.

C.3. PROYECTOS DE INVESTIGACIÓN

Este Apéndice recoge los principales proyectos europeos en el ámbito de la monitorización y teleasistencia de personas mayores. Para cada uno de ellos, la Tabla C.3 muestra su convocatoria, su duración y sus principales objetivos.

Tabla C.3. Principales proyectos de investigación en monitorización y teleasistencia

Título	Convocatoria	Objetivos
CONFIDENCE Ubiquitous care system to support independent living URL: http://www.confidence-eu.org/	FP7 2008-11	Sistema basado en sensores vestibles a través de los cuales se monitoriza el comportamiento del usuario. Se detectan eventos anómalos (p. e. caídas) e inesperados que puedan estar relacionados con problemas de la salud
Home Sweet Home Health monitoring and social integration environment for Supporting Wide Extension of independent life at HOME URL: http://www.homesweethome-project.be	CIP-ICT-PSP 2010-14	Soporte de las actividades de la vida diaria del usuario a través del uso de las TIC. El objetivo es la integración de dispositivos que permitan la monitorización y detección de alarmas. Entre ellos destacan sensores de variables médicas y sensores domóticos con los que detectar situaciones en las que el usuario necesite asistencia
Sensation-AAL Sensing and action to support mobility in Ambient Assisted Living URL: http://www.sensation-aal.eu/	AAL Call 1 2007-09	Asistencia de personas mayores para mantenerles independientes para las actividades de la vida diaria previniendo lesiones. Se basa en el uso de sensores vestibles para el reconocimiento y asistencia de actividades

Título	Convocatoria	Objetivos
HAPPY AGEING A Home based APPROach to the Years of AGEING URL: http://happyageing.info/	AAL Call 1 2008-11	Creación de un dispositivo para monitorizar el estilo de vida de las personas mayores. La idea es que el dispositivo sea capaz de monitorizar el comportamiento recordando la realización de actividades, detectando comportamiento inusuales y periodos de inactividad
Care Safe private homes for elderly persons URL: http://care-aal.eu/	AAL Call 1 2009-12	Reconocimiento automático y generación de alarmas ante la detección de situaciones críticas (p. e. caídas). El sistema se basa en el uso de un sensor de visión con procesamiento en tiempo real
IS-ACTIVE Inertial sensing systems for advanced chronic condition monitoring and risk prevention URL: http://www.is-active.eu	AAL Call 1 2009-12	Diseño de un sensor inercial que permita medir y analizar el comportamiento diario del usuario promoviendo la vida independiente
PAMAP Physical activity monitoring for ageing people URL: http://www.pamap.org	AAL Call 1 2009-12	Monitorización de la actividad física del usuario con el objetivo de facilitar su estancia fuera del hospital garantizando su rehabilitación. A través de una interfaz de televisión el usuario puede visualizar su estado y actividad
SOFTCARE URL: http://www.softcare-project.eu/	AAL Call 1 2009-12	Brazaletes con acelerómetro y comunicación inalámbrica integrada. Se realiza el reconocimiento de actividades y el envío de alarmas ante situaciones inusuales o de peligro para el usuario

Título	Convocatoria	Objetivos
AGNES Successful ageing in a networked society URL: http://agnes-aal.eu/site/	AAL Call 1 2009-12	A través de cámaras y teléfono móvil se realiza una monitorización del usuario detectando su estado de salud y las actividades que realiza. Permite comunicación entre usuario a través de una red social específica
eCAALYX Enhanced complete ambient assisted living experiment URL: http://www.ecaalyx.org	AAL Call 1 2009-12	Sistema completo de monitorización basado en el uso de sensores vestibles y ambientales. Se pretende realizar una monitorización global del usuario con el objetivo de mejorar su calidad de vida
H@H Health at home URL: http://www.health-at-home.eu/	AAL Call 1 2009-12	Comunicación entre usuarios y hospitales a través de sensores vestibles y arquitectura cliente-servidor. El sistema se enfoca en enfermedades cardiovasculares y respiratorias
REMPARK Personal Health Device for the Remote and Autonomous Management of Parkinson's Disease URL: http://www.rempark.eu/	FP7 2011-15	El objetivo es determinar y cuantificar los síntomas que presenta la enfermedad del Parkinson asociados a los estados motores. A partir de este conocimiento, el enfermo podrá ser supervisado de forma más efectiva mediante acciones personalizadas y ajustadas al usuario
Care@Home CARE services advancing the social interaction, health wellness and well-being of elderly people AT HOME URL: http://www.careathome-project.eu	AAL Call 3 2011-14	Monitorización remota y continua utilizando teléfonos móviles y sensores inalámbricos de emergencias en tiempo real

Título	Convocatoria	Objetivos
CaMeLi Care Me for Life URL: http://www.cameli.eu	AAL Call 5 2013-16	A través de sensores de visión se realiza una monitorización del comportamiento y estado anímico del usuario. Los cuidadores (tanto médicos como familiares) tienen acceso a la información recogida para su evaluación
Dalia Assistant for Daily Life Activities URL: http://www.dalia-aal.eu	AAL Call 5 2013-16	Uso de los sensores integrados en un teléfono móvil para medir las variables médicas de un usuario
HELICOPTER Healthy Life support through ComPrehensive Tracking of individual and Environmental Behaviors URL: http://www.helicopter-aal.eu/	AAL Call 5 2013-16	Monitorización cualitativa y cuantitativa a través de un conjunto de sensores heterogéneos del comportamiento diario del usuario. El sistema admite retroalimentación por parte del usuario y sus familiares para aconsejar y motivar su día a día
ELF@Home Elderly sELF-care based on sELF-check of health conditions and sELF-fitness at Home URL: http://www.elfathome.eu	AAL Call 5 2013-16	Creación de programas de ejercicios específicos para mayores. Cada usuario del sistema recibe un plan adecuado a su estado de salud. Para la monitorización de las variables médicas se utilizan dispositivo autónomos y sencillos
DOREMI Decrease of cOgnitive decline, malnutRition and sedEntariness by elderly empowerment in lifestyle Management and social Inclusion URL: http://www.doremi-fp7.eu/	FP7 2013-16	Detección y evaluación de los efectos de la malnutrición, sedentarismo y declive cognitivo en personas mayores. La solución desarrollada permitirá prolongar la capacidad funcional y cognitiva

C.4. MEDIDORES DE ACTIVIDAD FÍSICA

En este Apéndice se recoge información sobre varios dispositivos disponibles comercialmente que permiten la monitorización del comportamiento. La Tabla C.4 muestra el nombre e imagen de los principales dispositivos existentes, su descripción, características, tipo de comunicación e información acerca del tipo y duración de la batería.

Tabla C.4. Principales medidores comerciales de actividad física

Nombre	Características	Comunicación	Batería
Fitbit One⁹ <i>Dispositivo tipo "pinza"</i> 	Mide pasos, distancia, calorías consumidas, plantas subidas y calidad del sueño	Inalámbrica con PC/Móvil. Descarga automática.	Recargable. 5-7 días.
Fitbit Charge HR¹⁰ <i>Dispositivo en forma de pulsera que incorpora sensor de pulso cardíaco</i> 	Similar a Fitbit One pero permite la monitorización del pulso cardíaco	Inalámbrica con PC/Móvil. Descarga automática.	Recargable. Hasta 5 días.
Fitbit Zip¹¹ <i>Dispositivo tipo "pinza"</i> 	Mide pasos, distancia y calorías quemadas.	Inalámbrica con PC/Móvil. Descarga automática.	Pila de botón. 6 meses.

⁹ Fitbit One. URL: <http://www.fitbit.com/es/one>. Acceso: 15/02/2014.

¹⁰ Fitbit Charge HR. URL: <http://www.fitbit.com/es/chargehr>. Acceso: 02/02/2015.

¹¹ Fitbit Zip. URL: <http://www.fitbit.com/es/zip>. Acceso: 15/02/2014.

Nombre	Características	Comunicación	Batería
<p>Nike + Fuelband¹² <i>Dispositivo en forma de pulsera</i></p> 	<p>Mide nivel de actividad física, pasos y calorías consumidas.</p>	<p>Inalámbrica con Móvil y cableada con PC. Descarga mediante interacción del usuario.</p>	<p>Recargable. 1-4 días.</p>
<p>Jawbone Up¹³ <i>Dispositivo en forma de pulsera.</i></p> 	<p>Mide nivel de actividad física, pasos, calorías consumidas y calidad del sueño.</p>	<p>Cableada con Móvil. Descarga mediante interacción del usuario.</p>	<p>Recargable. 10 días.</p>
<p>Misfit Shine¹⁴ <i>Dispositivo integrable en varios complementos</i></p> 	<p>Mide calorías, calidad del sueño y reconoce actividades</p>	<p>Inalámbrica con Móvil. Descarga automática.</p>	<p>Pila de botón. 4 meses de duración.</p>
<p>Polar Loop¹⁵ <i>Dispositivo en forma de pulsera.</i></p> 	<p>Mide nivel de actividad física, pasos, calorías consumidas y calidad del sueño. Es sumergible y compatible con bandas torácicas.</p>	<p>Inalámbrica con PC/Móvil. Descarga automática.</p>	<p>Recargable. 5 días.</p>

¹² Nike + FuelBand. URL: http://www.nike.com/us/en_us/c/nikeplus-fuelband. Acceso: 15/02/2014.

¹³ Jawbone Up. URL: <https://jawbone.com/up>. Acceso: 15/02/2014.

¹⁴ Misfit Shine. URL: <http://misfit.com/products/shine>. Acceso: 03/02/2015.

Nombre	Características	Comunicación	Batería
Garmin Vivofit¹⁶ <i>Dispositivo en forma de pulsera</i> 	Mide nivel de actividad física, pasos, calorías consumidas, pulso y calidad del sueño. Es sumergible y compatible con otro tipo de dispositivos de la marca.	Inalámbrica con PC/Móvil. Descarga mediante interacción del usuario.	No recargable. 1 año de duración.
Withings Pulse¹⁷ <i>Dispositivo integrable en pulsera pero que puede ser portado de otras formas.</i> 	Mide pasos, calorías consumidas, altitud, pulso cardíaco y calidad del sueño.	Inalámbrica con Móvil. Descarga automática.	Recargable. 15 días de duración.
Microsoft Band¹⁸ <i>Dispositivo en forma de pulsera</i> 	Mide pulso cardíaco, distancia recorrida, pasos, calorías y calidad del sueño	Inalámbrica con Móvil. Descarga automática.	Recargable. 2 días de duración.

¹⁵ Polar Loop. URL: http://www.polar.com/es/productos/get_active/fitness_crosstraining/loop. Acceso: 06/09/2014.

¹⁶ Garmin VivoFit. URL: <http://sites.garmin.com/vivo/>. Acceso: 06/09/2014.

¹⁷ Withings Pulse. URL: <http://www.withings.com/es/withings-pulse.html>. Acceso: 06/09/2014.

¹⁸ Microsoft Band. URL: <http://www.microsoft.com/Microsoft-Band/en-us>. Acceso: 03/02/2015.

Nombre	Características	Comunicación	Batería
<p>Sony Smartband¹⁹ <i>Dispositivo en forma de pulsera pero desmontable</i></p> 	<p>Registra automáticamente datos referentes a actividades físicas, sociales y lúdicas</p>	<p>Inalámbrica con Móvil. Descarga mediante interacción del usuario.</p>	<p>Recargable. 4-5 días de duración.</p>
<p>Samsung Gear Fit²⁰ <i>Dispositivo en forma de pulsera</i></p> 	<p>Se enfoca como un híbrido entre pulsera cuantificadora y reloj inteligente. Mide actividad, calorías, pasos...</p>	<p>Inalámbrica con móvil. Descarga automática.</p>	<p>Recargable. 2-3 días de duración.</p>
<p>Runtastic Orbit²¹ <i>Dispositivo en forma de pulsera o "pinza"</i></p> 	<p>Registra pasos, distancia, calorías, sueño y es sumergible.</p>	<p>Inalámbrica con móvil. Descarga manual o automática.</p>	<p>Recargable. 5-6 días de duración.</p>

¹⁹ Sony Smartband. URL: <http://www.sonymobile.com/es/products/smartwear/smartband-swr10/>. Acceso: 03/02/2015.

²⁰ Samsung Gear Fit. URL: <http://www.samsung.com/es/consumer/mobile-phone/wearables/wearables/SM-R3500ZKAPHE>. Acceso: 03/03/2015.

²¹ Runtastic Orbit. URL: <https://www.runtastic.com/es/orbit>. Acceso: 03/02/2015.

APÉNDICE D.

ESTUDIO DEL MÉTODO DE MEDIDA

Este apéndice recoge la información de los 39 usuarios participantes en el experimento realizado para la validación de JIM como estimador de la actividad física. Este experimento es descrito en la sección 4.4.4 del Capítulo 4. Cada usuario fue monitorizado durante al menos 24 horas.

Para cada usuario, la Tabla D.1 muestra su sexo, edad, peso, estatura e índice de masa corporal (BMI). Para la identificación del perfil de actividad se ha tenido en cuenta el tipo de ejercicio físico realizado durante el periodo de monitorización (duración experimento). La Tabla D.2 muestra los datos obtenidos durante la experimentación: índice de correlación, JIM_{total} y SMA_{total} . El índice de correlación es uno de los parámetros más importantes ya que representa la fuerte asociación existente entre los dos estimadores estudiados: JIM y SMA.

Tabla D.1 Lista de usuarios involucrados en la validación del método de medida

Nº	Sexo	Edad	Peso (Kg)	Estatura (m)	BMI	Duración experimento (min)	Perfil de actividad
1	M	27	57,0	1,70	19,72	1859	Medio
2	M	28	76,0	1,63	28,60	1823	Bajo
3	M	25	45,0	1,56	18,49	1788	Bajo
4	M	27	69,0	1,69	24,16	1833	Bajo
5	M	28	61,0	1,76	19,69	1804	Medio
6	V	31	80,0	1,81	24,42	1916	Medio
7	V	32	73,5	1,77	23,46	1916	Alto
8	V	33	66,0	1,70	22,84	1916	Medio
9	M	28	55,0	1,59	21,76	1916	Alto
10	V	29	78,0	1,82	23,55	1555	Alto
11	V	29	67,0	1,74	22,13	1591	Alto
12	M	28	57,0	1,69	19,96	1653	Alto
13	V	27	81,0	1,78	25,56	1641	Alto
14	V	29	80	1,75	26,12	1663	Bajo
15	M	38	60	1,59	23,73	1916	Medio
16	M	57	69	1,58	27,64	1916	Bajo
17	M	47	72	1,65	26,45	1913	Bajo
18	V	54	75	1,64	27,89	1909	Medio
19	M	41	45	1,60	17,58	1916	Medio
20	V	41	109	1,64	40,53	1916	Medio
21	M	54	62	1,60	24,22	1913	Bajo
22	M	49	58	1,58	23,23	1909	Medio
23	M	50	57	1,61	21,99	1913	Bajo
24	M	46	69	1,69	24,16	1913	Medio
25	V	57	80	1,75	26,12	1853	Medio
26	M	50	68	1,63	25,59	1909	Medio
27	M	56	64	1,58	25,64	1677	Medio
28	V	28	88	1,78	27,77	1522	Alto
29	M	26	63	1,60	24,61	1838	Bajo
30	M	29	64	1,64	23,80	1724	Bajo
31	V	26	61	1,75	19,92	1761	Alto
32	V	29	83	1,86	23,99	1913	Alto
33	V	58	66	1,67	23,67	1913	Bajo
34	M	51	57	1,54	24,03	1916	Medio
35	V	27	73	1,78	23,04	1916	Alto
36	V	38	82	1,69	28,71	1916	Bajo
37	V	31	88	1,76	28,41	1511	Alto
38	V	32	70	1,73	23,39	1618	Medio
39	M	45	90	1,70	31,14	1606	Medio

Tabla D.2 Resultados de la validación del método de medida

Nº	Índice de correlación JIM-SMA	JIM _{total}	SMA _{total}	Diferencia media JIM-SMA
1	0,99	23915	18196	3,08
2	0,96	20460	15952	2,47
3	0,95	20242	15993	2,38
4	0,97	23548	18090	2,98
5	0,97	19441	14692	2,63
6	0,98	25300	18949	3,31
7	0,92	21611	16751	2,54
8	0,94	21930	17107	2,52
9	0,97	27643	20890	3,52
10	0,87	19785	14301	3,53
11	0,96	21462	16629	3,04
12	0,97	22896	17376	3,34
13	0,98	20266	15623	2,83
14	0,96	16290	12824	2,08
15	0,96	24895	19168	2,99
16	0,94	23026	18166	2,54
17	0,97	24150	18675	2,86
18	0,95	20813	16287	2,37
19	0,90	19781	16073	1,94
20	0,91	25429	20178	2,74
21	0,93	22214	17372	2,53
22	0,97	28975	22197	3,55
23	0,93	22150	17217	2,58
24	0,92	22035	17050	2,61
25	0,95	17841	13893	2,13
26	0,98	27353	21033	3,31
27	0,98	27166	20528	3,96
28	0,89	21756	16067	3,74
29	0,96	31023	23665	4,00
30	0,95	14592	11384	1,86
31	0,96	23925	18360	3,16
32	0,98	28298	21395	3,61
33	0,89	19441	15696	1,96
34	0,92	26832	22517	2,25
35	0,98	26794	21086	2,98
36	0,99	23859	17991	3,06
37	0,86	15422	11838	2,37
38	0,96	12560	10339	1,37
39	0,97	26467	20754	3,56

APÉNDICE E.

ESTUDIO DEL MÉTODO DE MODELADO

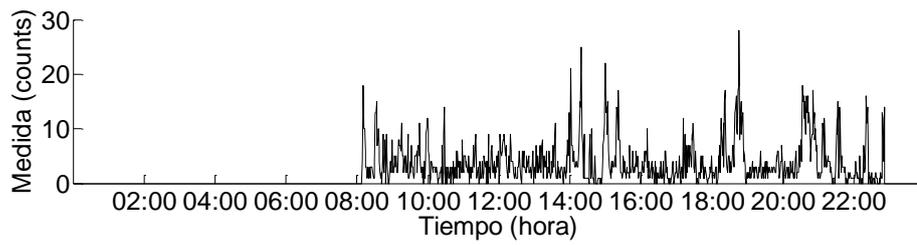
E.1. ANÁLISIS DE LAS PRINCIPALES TÉCNICAS DE SEGMENTACIÓN

E.1.1. INTRODUCCIÓN

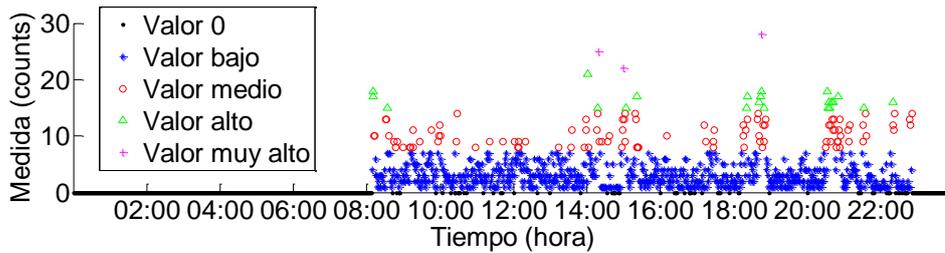
En este apéndice se analiza la aplicación de las principales técnicas de segmentación sobre la variable de medida de la actividad física. En primer lugar se analiza la aplicación de las técnicas de segmentación por regiones (sección E.1.2), a continuación las técnicas de segmentación por detección de bordes (sección E.1.3) y en último lugar las técnicas de inventariado temporal (sección E.1.4). Todas estas técnicas fueron documentadas en la sección 5.2.2.2 del Capítulo 5.

E.1.2. SEGMENTACIÓN POR REGIONES

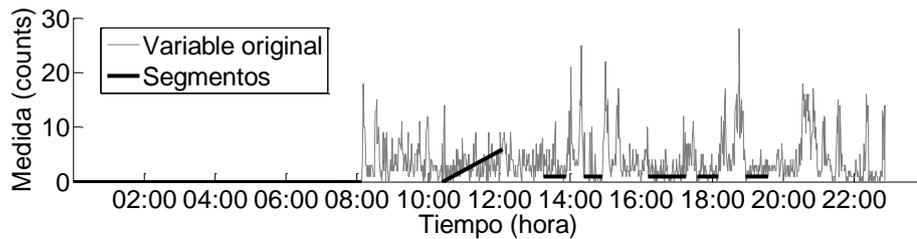
Dentro de las técnicas de segmentación por regiones destacan las técnicas de umbralización (Figura E.1.b), las técnicas de detección de patrones (Figura E.1.c) y los algoritmos *top-down* (Figura E.1.d y Figura E.1.e) tal y como se mencionó en la Tabla 5.5. A continuación se describe la aplicación de cada una de estas técnicas y los resultados obtenidos sobre una señal de ejemplo mostrada en la Figura E.1.a.



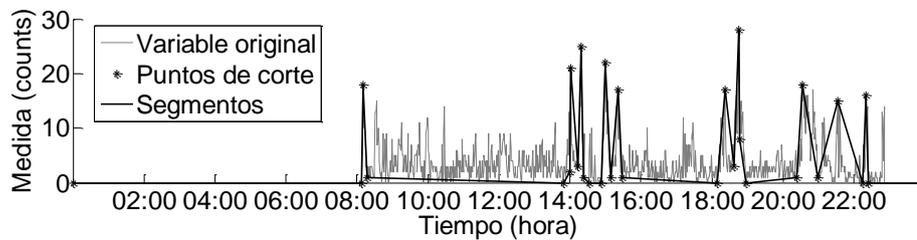
a) Variable original



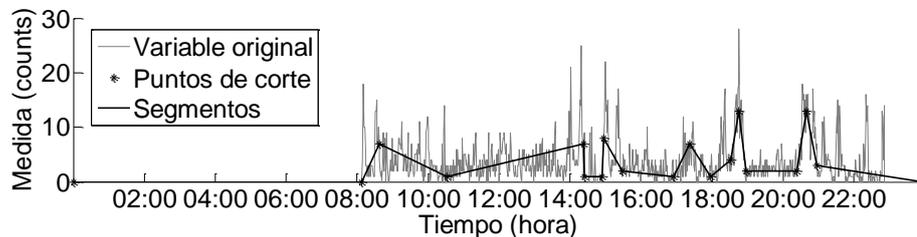
b) Segmentación por regiones: umbralización



c) Segmentación por regiones: transformada de Hough



d) Segmentación por regiones: aproximación poligonal sobre variable original



e) Segmentación por regiones: aproximación poligonal sobre variable filtrada

Figura E.1. Aplicación de técnicas de segmentación por regiones

Las técnicas de segmentación por umbralización se basan en la división de la variable en función de su nivel de medida. El resultado será un conjunto de segmentos cuyas muestras pertenecen al mismo nivel de intensidad. En el ejemplo de la Figura E.1.b la segmentación se realiza de forma proporcional al máximo valor medido, de forma que se obtienen cinco grupos de muestras:

- $Medida \leq 0$; puntos en color negro: "Valor 0".
- $0 < Medida \leq \text{Máximo}/4$; asteriscos en color azul: "Nivel bajo".
- $\text{Máximo}/4 < Medida \leq 2*\text{Máximo}/4$; círculos en color rojo: "Nivel medio".
- $2*\text{Máximo}/4 < Medida \leq 3*\text{Máximo}/4$; triángulos en color verde: "Nivel alto".
- $Medida \leq \text{Máximo}$; cruces en color magenta: "Nivel muy alto".

Para realizar la segmentación por umbralización se pueden implementar diferentes aproximaciones fijando los umbrales estáticamente o dinámicamente en función de las características de la variable a segmentar. Uno de los principales inconvenientes en la aplicación de técnicas de umbralización es la selección de estos valores. Este inconveniente puede ser resuelto implementando técnicas de *clustering* u otro tipo de técnicas de aprendizaje automático.

Como ejemplo de técnica de segmentación por detección de patrones se ha aplicado la transformada de Hough sobre la variable de la Figura E.1.a obteniendo las líneas negras representadas en la Figura E.1.c. La transformada de Hough permite la detección de patrones en variables y aunque su principal aplicación es la detección de formas en imágenes, se ha realizado una adaptación de la implementación de la transformada para la detección de líneas en variables unidimensionales como la actividad física. Los parámetros en la detección han sido fijados para la detección de líneas rectas perpendiculares al eje y ($\theta \in [85^\circ, 90^\circ]$) y que se corresponden con segmentos constantes de la variable. Como longitud mínima de la línea detectada se ha fijado 30 muestras.

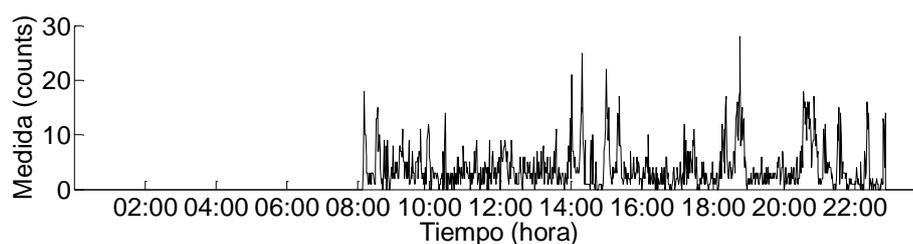
Esta aproximación para la segmentación de una variable unidimensional continua ha sido utilizada en trabajos como (Molleda et al. 2012). Uno de los principales inconvenientes del uso de esta transformada es que se requiere de una gran parametrización y de conocimiento experto acerca de la variable sobre la que se realiza la segmentación. Cambios en la variable pueden hacer que los parámetros seleccionados dejen de ser válidos. Sin embargo, si el patrón que sigue la variable es conocido, el uso de esta transformada permite realizar una segmentación óptima por regiones y un procesamiento independiente para cada región detectada.

En la Figura E.1.d se muestra la aplicación de un algoritmo *top-down* sobre la actividad física utilizada como ejemplo (Figura E.2.a). Esta técnica se basa en la aproximación recursiva de una variable por un conjunto de líneas rectas intentando minimizar el error que se comete en el ajuste. El error permitido es uno de los parámetros de entrada al algoritmo. En la Figura E.1.d se muestra la aplicación del algoritmo directamente sobre la variable aceptando un error de 4 veces su desviación estándar. Debido al ruido existente en la variable y a las variaciones repentinas que se producen en ella (ruido impulsivo), la aproximación obtenida descompone la variable en un gran número de segmentos. Cada asterisco representa cada uno de los puntos determinados por el algoritmo y el conjunto de segmentos negros constituye la aproximación lineal, resultado del algoritmo.

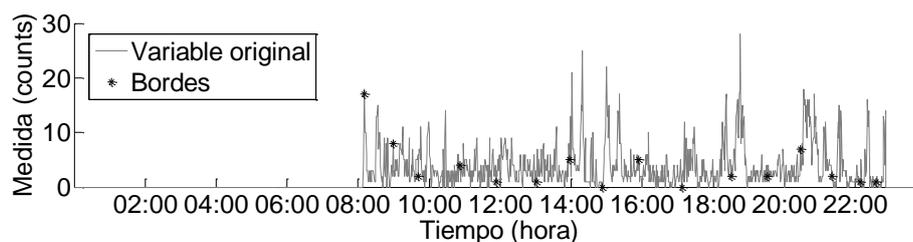
Para mejorar los resultados del algoritmo *top-down* una posible opción es aplicar un filtrado de mediana para eliminar el ruido impulsivo y suavizar la variable sin que pierda su forma original. El filtro de mediana aplicado tiene un tamaño de 15 muestras y hace que el conjunto de segmentos obtenido (Figura E.1.e) caracterice de una mejor forma a la variable original. Sin embargo, en función del error permitido en el ajuste (parámetro de entrada), el resultado de la segmentación variará. Si el error permitido es muy grande, el número de segmentos será muy bajo y caracterizará vagamente a la variable original. Si el error permitido es muy bajo, el número de segmentos podrá llegar a ser muy alto y la caracterización de la variable será similar a trabajar con la variable sin segmentar. El uso de este tipo de algoritmos para la segmentación implica alcanzar un compromiso entre ambos extremos mediante la selección de un parámetro de entrada adecuado.

E.1.3. SEGMENTACIÓN POR DETECCIÓN DE BORDES

Dentro de las técnicas de detección de bordes, la Figura E.2.b muestra la aplicación del algoritmo de Canny sobre la variable de la Figura E.2.a (misma variable de ejemplo que la utilizada en la Figura E.1). Cada uno de los asteriscos negros representados sobre la variable original muestra los bordes identificados por el algoritmo. Estos resultados podrían ser optimizados mediante un análisis escala-espacio mejorando la detección y maximizando la precisión en la localización. La precisión indicada en el algoritmo de detección de bordes aplicado es de 15 muestras.



a) Variable original



b) Segmentación por fronteras: algoritmo de Canny

Figura E.2. Aplicación de técnicas de segmentación por fronteras

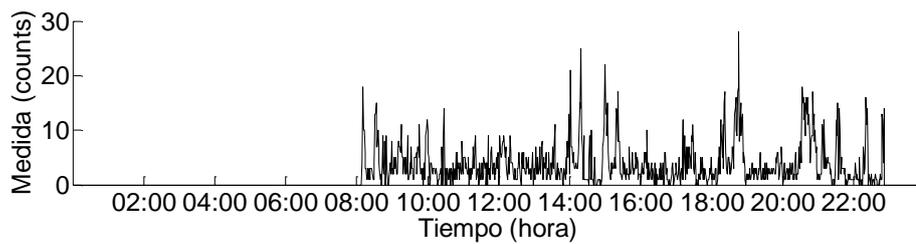
E.1.4. ENVENTANADO TEMPORAL

En general, las técnicas de detección de regiones y de detección de fronteras son útiles cuando se va a realizar un procesamiento diferente en función del nivel o de las características de la variable medida. Así, por ejemplo, la variable de salida representada en la Figura E.1.e permitiría obtener una división de la actividad en función del patrón que sigue y realizar un procesamiento independiente en función de si detecta un patrón de actividad ascendente, descendente o constante.

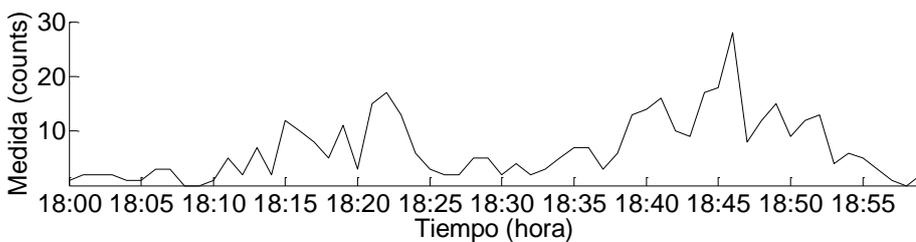
En el apartado 5.2.2.2 se mostraron tres tipos diferentes de ventanas que podrían ser aplicadas sobre la actividad física para su segmentación: ventana cuadrada, ventana triangular y ventana de Hamming. La Figura E.3 muestra la aplicación de las tres técnicas de enventanado sobre la misma variable que se utilizó en la Figura E.1 y Figura E.2 y que vuelve a representarse en la Figura E.3.a. Partiendo de un tamaño de ventana prefijado ($N_m = 60$) se muestra el enventanado para cada tipo de ventana en la muestra $n = 1110$ correspondiente con la actividad física medida a las 18:30 horas.

El uso de una ventana triangular (ver Figura E.3.c) o una ventana de Hamming (ver Figura E.3.d) hace que la segmentación de la variable tenga lugar simultáneamente a una transformación en los valores segmentados. Este tipo de transformaciones son interesantes cuando se desea dar más peso a la muestra central de la ventana durante el procesamiento y realizar simultáneamente una transformación en las adyacentes.

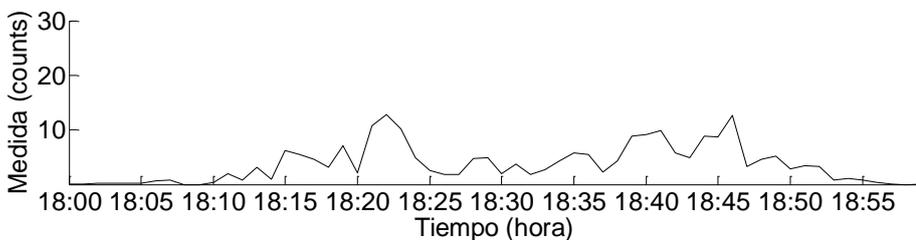
Al contrario que en los casos anteriores, la ventana cuadrada (Figura E.3.b) no realiza ninguna transformación sobre los datos de entrada. Este tipo de ventana es útil cuando lo único que se desea es realizar una división de la señal de entrada en segmentos de menor tamaño para facilitar su procesamiento. Puesto que en este trabajo, ese es el objetivo de la etapa de segmentación, una ventana cuadrada resulta adecuada. En su selección es necesario definir el tamaño de la misma y el número de muestras de solapamiento.



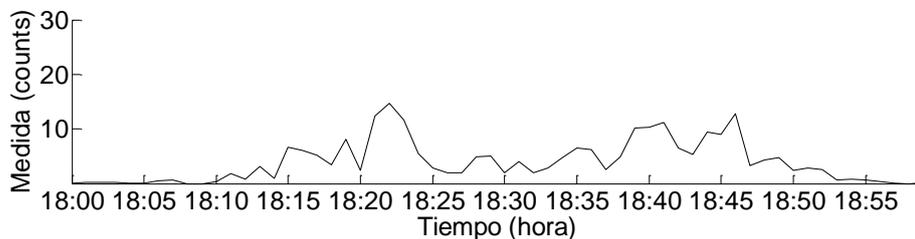
a) Variable original



b) Aplicación de ventana cuadrada



c) Aplicación de ventana triangular



d) Aplicación de ventana de Hamming

Figura E.3. Aplicación de técnicas de enventanado temporal

E.2. LISTA DE USUARIOS

Esta sección recoge la información de los 8 usuarios participantes en el experimento realizado para la validación del método de modelado y la identificación de la rutina del usuario en el Capítulo 5. Cada usuario fue monitorizado durante 2 días no consecutivos. Los resultados obtenidos son descritos y analizados en la sección 5.3.4.

La Tabla E.1 muestra para cada uno de los usuarios involucrados en la experimentación su sexo, edad, peso, estatura e índice de masa corporal (BMI). El perfil de actividad del usuario está relacionado con el tipo de ejercicio físico realizado durante el periodo de monitorización. Los usuarios con perfil de actividad bajo no realizaron ningún tipo de ejercicio físico y los usuarios con perfil de actividad alto llevaron a cabo actividades deportivas de alta intensidad. El índice de correlación muestra la asociación existente entre la rutina identificada mediante el método de modelado propuesto y la rutina teórica indicada por los usuarios. Los valores obtenidos muestran la fuerte asociación existente entre ambos valores.

Tabla E.1. Lista de usuarios y resultados de la validación del método de modelado

Nº	Sexo	Edad	Peso (Kg)	Estatura (m)	BMI	Perfil de actividad	Índice de correlación rutina
1	V	64	77,0	1,76	24,86	Medio	0,94
2	M	63	54,0	1,72	18,25	Bajo	0,93
3	V	71	75,0	1,60	29,29	Bajo	0,95
4	V	78	70,0	1,65	25,71	Medio	0,93
5	V	31	64,0	1,70	22,10	Medio	0,91
6	M	31	80,0	1,62	30,48	Medio	0,90
7	M	37	54,0	1,60	21,09	Medio	0,92
8	M	36	52,0	1,59	20,57	Alto	0,87

APÉNDICE F.

ESTUDIO DEL MÉTODO DE DETECCIÓN

F.1. ESTUDIO DE MÉTODOS SEMI-SUPERVISADOS

F.1.1. INTRODUCCIÓN

Dentro de los métodos semi-supervisados, tal y como se mencionó en la sección 6.2.4 destacan los métodos de ajuste a una distribución conocida, el cálculo del histograma, los diagramas de cajas y el uso de medidas de similitud. A continuación, se valora el uso de cada uno de estos métodos para la detección de anomalías en la actividad física.

F.1.2. AJUSTE A UNA DISTRIBUCIÓN NORMAL

Dentro de los métodos paramétricos para la detección de anomalías se encuentran las técnicas que asumen que los datos siguen una distribución normal de parámetros μ (media) y σ (desviación estándar). La Figura F.1 muestra la función de densidad de probabilidad de una distribución normal en función de sus parámetros característicos μ (media) y σ (desviación estándar). La expresión matemática de la función de densidad de probabilidad de este tipo de distribución se muestra en la Ecuación F.1.

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(x-\mu)^2}{2 \cdot \sigma^2}}$$

Ecuación F.1

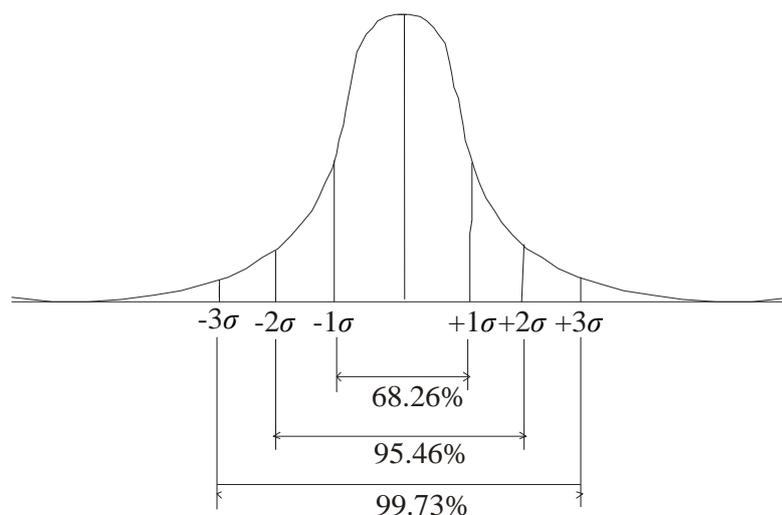


Figura F.1. Función de densidad de probabilidad de una distribución normal

Basándose en el ajuste de las muestras a una distribución de este tipo, la detección de anomalías tiene lugar mediante el estudio de la distancia de una muestra a la media de la distribución. Este valor puede ser utilizado directamente como puntuación de anomalía ya que identifica lo lejos que una muestra de actividad se encuentra de la media de la distribución (valor con mayor probabilidad de aparición).

La técnica más utilizada en el estado del arte tras el ajuste de los datos a una distribución normal es la conocida como la regla 3σ . Esta aproximación se basa en la identificación como anomalías de todas las muestras de una variable que se encuentran a una distancia mayor de 3σ de la media de las medidas. Esta aproximación parte de la asunción de que en una distribución normal el 99,73% de las medidas se encuentran en la región $\mu \pm 3\sigma$ tal y como se representa en la Figura F.1. Existen otro tipo de técnicas que pueden ser utilizadas para la detección de anomalías como el test de Grubbs (Grubbs 1969) o el test de Dixon (Dean and Dixon 1951). Todas estas técnicas parten de la asunción de que los datos de medida siguen una distribución normal.

Para valorar la aplicación de la regla 3σ a la detección de anomalías en la actividad física, es necesario determinar si es posible el ajuste de las medidas a una distribución normal. Para ello se parten de las medidas de ejemplo descritas en el Apéndice H y que se corresponden con la monitorización de un usuario durante cinco días.

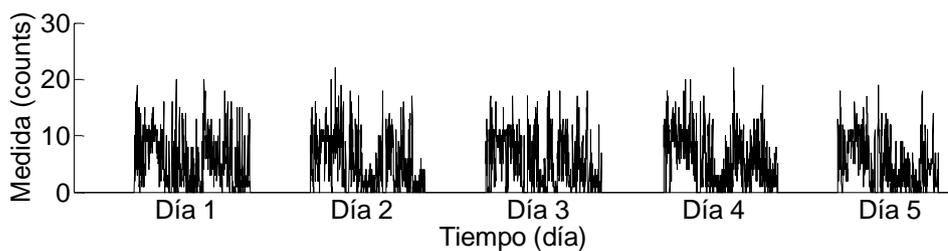
La Figura F.2 muestra el análisis realizado sobre las medidas de actividad y sus intentos de ajuste a una distribución normal. Partiendo del conjunto de medidas sobre los cinco días de duración del experimento (Figura F.2.a) se construye el diagrama de densidad de probabilidad de la Figura F.2.b sobre el que se muestra el ajuste a una

distribución normal de media $\mu = 3,41$ y desviación $\sigma = 4,28$. Aunque gráficamente ya se aprecia como la distribución de medidas no se ajusta adecuadamente, existen diferentes test estadísticos que pueden ser aplicados sobre los datos para evaluar su ajuste a una distribución normal. La aplicación del Test Lilliefords sobre los datos ajustados determina que los datos de medida no siguen una distribución normal.

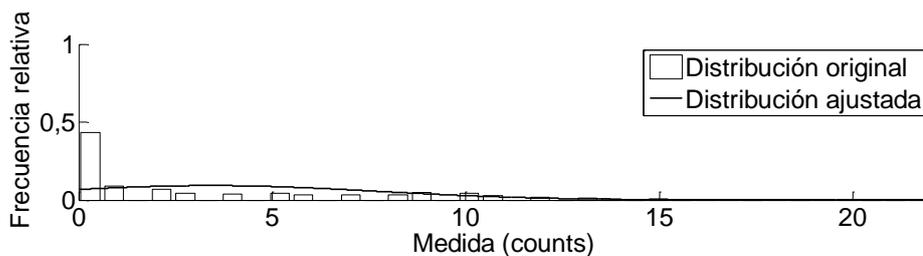
Una aproximación común, cuando los datos originales de medida no siguen una distribución normal, es la transformación de los mismos para conseguir que los datos transformados sí se ajusten adecuadamente. En el estado del arte destacan principalmente tres transformaciones que partiendo de los datos originales x pueden conseguir que los datos transformados se ajusten a una distribución normal: $\ln(x + 1)$, $\sqrt{x + 1}$ y $1/(x + 1)$. La Figura F.2.c, Figura F.2.d y Figura F.2.e muestra cada una de estas transformaciones y su ajuste a una distribución normal. La aplicación del test Lilliefords confirma que ninguna de las transformaciones consigue el ajuste de los datos a una distribución normal.

Otra aproximación para el ajuste de los datos a una distribución normal es el uso de medidas de actividad parciales en lugar de las medidas totales. Es decir, que el ajuste a una distribución normal se lleve a cabo tras la segmentación de la variable. De esta forma, el ajuste se estudiará para una determinada muestra n en los segmentos de medidas de actividad extraídos para ese instante temporal en los N_r días de actividad. Cada segmento de actividad cubrirá los instantes comprendidos entre $n - N_m/2$ y $n + N_m/2$ para cada día de monitorización. En el caso de que las medidas de los segmentos si se ajusten a una distribución normal, la detección de anomalías se produciría de forma adaptativa para cada instante temporal mediante el ajuste a una distribución de parámetros diferentes.

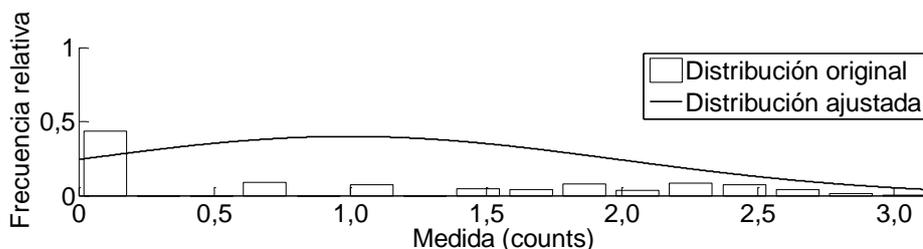
El estudio del ajuste de las muestras anteriormente descritas (segmentos para cada muestra n de $n - N_m/2$ a $n + N_m/2$ en los $N_r = 5$ días anteriores) ha sido estudiado de forma analítica. Igual que ocurría con las medidas totales de actividad, las medidas parciales para cada segmento tampoco se ajustan a una distribución normal ni utilizando los datos originales, ni las transformaciones analizadas en la Figura F.2. En ninguno de los casos, el test Lilliefords ha confirmado que los datos de medida sigan una distribución normal.



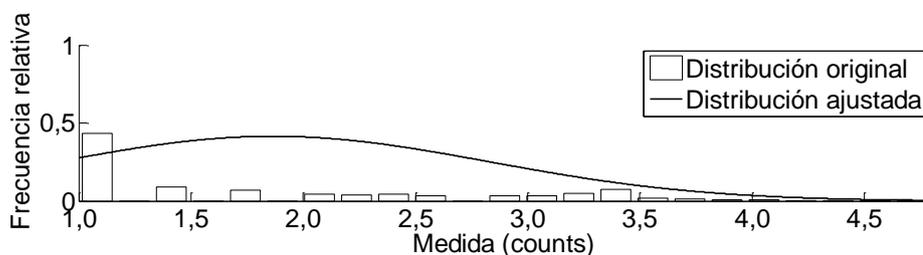
a) Medidas de actividad durante los cinco días de duración del experimento



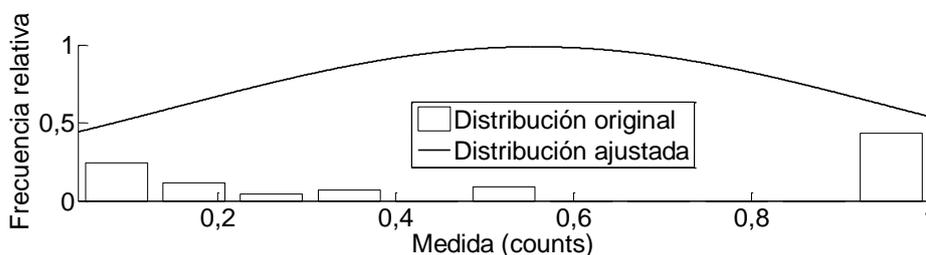
b) Densidad de probabilidad y ajuste a una distribución normal



c) Densidad de probabilidad con transformación $\ln(x+1)$ y ajuste a una distribución normal



d) Densidad de probabilidad con transformación $\sqrt{x+1}$ y ajuste a una distribución normal



e) Densidad de probabilidad con transformación $1/(x+1)$ y ajuste a una distribución normal

Figura F.2. Ajuste a una distribución normal de las medidas de actividad

En vista del estudio realizado, el ajuste a una distribución normal no se trata de una técnica adecuada para la detección de anomalías en el comportamiento humano. A su vez, el hecho de asumir que los datos siguen una distribución normal puede conllevar errores cuando en un determinado momento los datos dejen de seguirla o cuando aparezcan nuevas rutinas de comportamiento. También, al trabajar con múltiples usuarios es complicado confirmar que una determinada distribución es seguida para cada uno de los usuarios monitorizados.

F.1.3. CÁLCULO DE HISTOGRAMAS

Dentro de los métodos estadísticos no paramétricos para la detección de anomalías destacan las técnicas basadas en el cálculo de histogramas. Este tipo de técnicas no asumen la existencia de ninguna distribución conocida en los datos de medida y se basan en el cálculo del histograma para los datos de entrada considerados como normales. Están formados por dos fases: en primer lugar se calcula el histograma sobre un conjunto de datos de entrada y a continuación se realiza la detección de anomalías estudiando la relación entre cada muestra y el histograma calculado.

En el uso de histogramas para la detección de anomalías hay dos estrategias diferenciadas:

- Considerar como anomalía toda aquella muestra que no se encuentre dentro de los rangos del histograma calculado
- Puntuar la anomalía de cada muestra en función de su frecuencia de aparición en el rango del histograma calculado

La segunda estrategia resulta mucho más interesante en el campo del comportamiento humano ya que permite tener para cada muestra de actividad una puntuación asociada con su anomalía. Este tipo de métodos identifican como anómalas a aquellas muestras con baja frecuencia de aparición. Se trata de una técnica muy sencilla cuando se trabaja con datos unidimensionales pero que se complica considerablemente cuando se trabaja con colecciones de atributos.

La puntuación de anomalía puede ser calculada directamente como la frecuencia de aparición del rango en el histograma o mediante la relación con la frecuencia máxima de aparición. Si se trabajara con una distribución de datos uniforme, la primera de las estrategias proporcionaría valores de anomalía para todas las muestras mientras que la segunda de las estrategias detectaría normalidad para todas ellas. Por esta razón, resulta mucho más interesante puntuar la anomalía de una muestra en función de su

frecuencia de aparición relativa a la moda de la distribución (rango con mayor frecuencia de aparición).

Una de las claves en el uso de histogramas para la detección de anomalías es la selección del número de rangos de medida utilizados para la construcción del histograma (Chandola et al. 2009). Si se trabaja con un valor bajo para el cálculo de los rangos, muchas muestras normales pueden verse dentro de rangos vacíos o extraños. Por el contrario, si se trabaja con un valor alto, muchas muestras anómalas podría verse dentro de rangos del histograma con gran frecuencia de aparición y el método utilizado presentaría una alta tasa de falsos negativos. Para resolver este problema y llegar a un compromiso entre ambas situaciones, se propone el uso de la regla Freedman–Diaconis (Freedman and Diaconis 1981) que calcula el tamaño óptimo de los rangos a utilizar en el cálculo de un histograma (ver Ecuación F.2). El parámetro IQR representa la distancia inter-cuartil en la serie x sobre la que se calcula el histograma y N el número de observaciones de la serie.

$$\text{Tamaño rango} = 2 \cdot \text{IQR}(x) \cdot N^{-1/3}$$

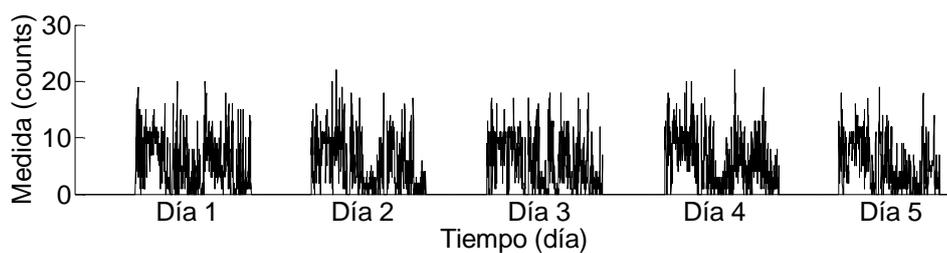
Ecuación F.2

Partiendo de los mismos datos de experimentación que se utilizaron en la sección F.1.2 (datos descritos en el Apéndice H), la Figura F.3 representa el cálculo del histograma para los datos medidos en un usuario de ejemplo durante cinco días. Para calcular los rangos del histograma se utilizó la regla Freedman-Diaconis anteriormente descrita (ver Ecuación F.2). A partir de estos datos se construye el histograma de normalidad que se muestra en la Figura F.3.c y que permite conocer la frecuencia relativa de aparición de una determinada medida de actividad. Un valor alto de histograma (próximo a 1) indica una alta frecuencia de aparición y un valor bajo (próximo a 0), una baja frecuencia relativa de aparición.

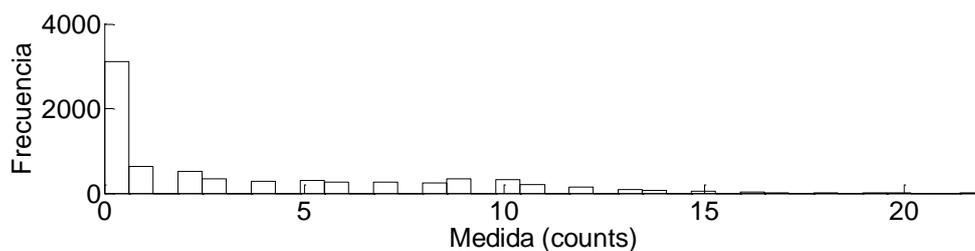
Usando la aproximación descrita, la Figura F.3.e muestra la aplicación del algoritmo de detección de anomalías sobre la medida del día evaluable representada en la Figura F.3.d. La salida del algoritmo es una variable entre 0 y 1 donde 0 representa normalidad y los valores cercanos a 1 la anomalía. Para ello, partiendo del valor del histograma normalizado (H_n) la puntuación de anomalía se calcula siguiendo la expresión de la Ecuación F.3.

$$\text{Puntuación} = \text{abs}(H_n(x) - 1)$$

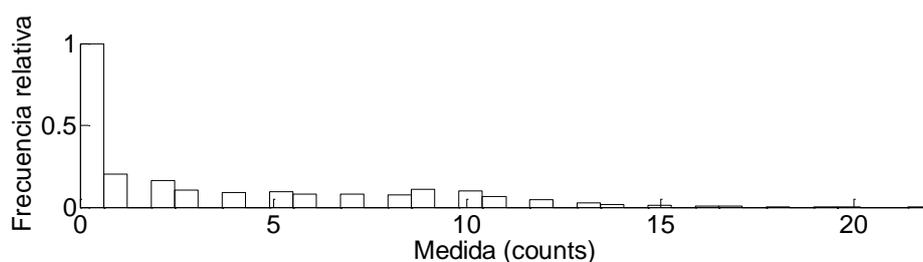
Ecuación F.3



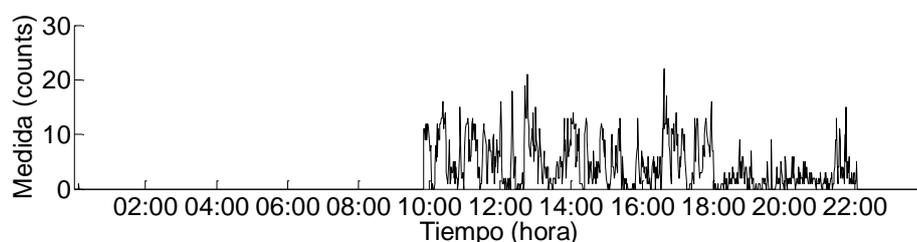
a) Medidas de actividad durante los cinco días de duración del experimento



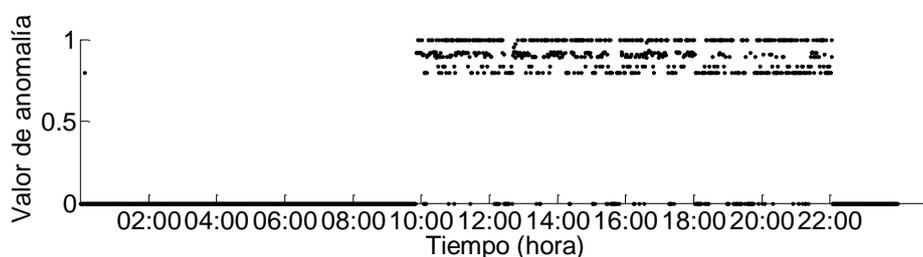
b) Histograma de las medidas de actividad



c) Histograma normalizado de las medidas de actividad



d) Medida de actividad sobre un día evaluable



e) Detección de anomalías mediante el histograma normalizado

Figura F.3. Cálculo de histograma para detección de anomalías

El uso de este método basado en el cálculo del histograma para la detección de anomalías presenta dos inconvenientes principales:

- La anomalía detectada es independiente del contexto. Al estar trabajando con el histograma sobre todo el conjunto de medidas de actividad durante los $N_r = 5$ días, la puntuación de anomalía obtenida es independiente del instante temporal de medida y solo depende de la frecuencia de aparición en el conjunto total.
- Los valores altos de actividad son puntuados siempre con un valor alto de anomalía. Esto es debido a que generalmente el usuario se encontrará un mayor número de horas en reposo que con niveles de actividad altos. Al trabajar sobre el conjunto total de medidas de actividad, las muestras correspondientes con las horas de sueño siempre supondrán un porcentaje alto respecto al resto de medidas. Así, cualquier valor alto de actividad se detectará como anómalo con una puntuación alta.

Estos inconvenientes se ponen de manifiesto en el análisis de la Figura F.3.d donde el método utilizado siempre valora con puntuación 0 (normalidad) a las muestras correspondientes con el reposo y con valor 1 (anomalía máxima) a las muestras correspondientes con un nivel de actividad alto. En el tramo horario entre las 08:00 y las 10:00 (ver Apéndice H) el método de detección no es capaz de identificar la anomalía que tiene lugar cuando el usuario duerme dos horas más de lo habitual ya que el reposo siempre es valorado como normalidad independientemente del contexto de medida. A su vez, si aparecieran nuevos valores de actividad, estos siempre serán puntuados con un valor 1 de anomalía ya que su frecuencia de aparición en los datos utilizados para la construcción del histograma es nula.

Debido a estos inconvenientes, esta primera aproximación para el uso de histogramas en la detección de anomalías no es adecuada. Siguiendo una aproximación similar a la llevada a cabo en la sección F.1.2, los histogramas podrían ser construidos para cada uno de los segmentos calculados sobre los $N_r = 5$ días para un mismo instante temporal. Cada uno de estos histogramas será normalizado tal y como se mostró en la Figura F.3.b y se utilizará la expresión de la Ecuación F.3 para obtener la puntuación de anomalía.

La Figura F.4 muestra la detección de anomalías utilizando la técnica descrita sobre el mismo ejemplo de la Figura F.3. En este caso, el método de detección de anomalías es capaz de detectar el cambio de comportamiento que tiene lugar durante la actividad matutina del usuario al contrario de lo que ocurría en la detección de la Figura F.3.e. Se detectan anomalías entre las 08:30 y las 10:00 referente al cambio de hora de

despertarse y al descenso de actividad. Parece por tanto, a priori, que un método de detección de anomalías basado en el cálculo del histograma para cada segmento de actividad puede ser una opción válida para el análisis del comportamiento. Su combinación con un método de unión de puntos permitiría la identificación de periodos de tiempo de comportamiento anómalo.

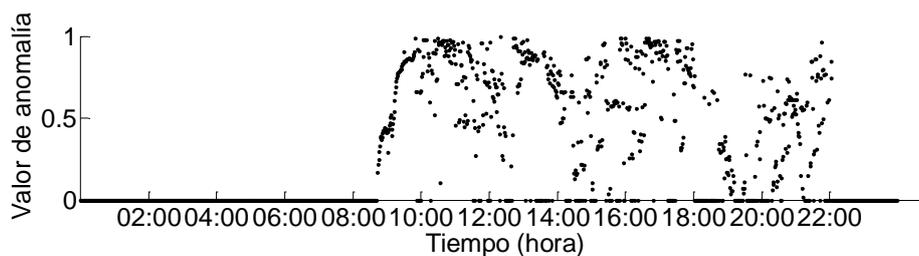


Figura F.4. Detección de anomalías mediante histograma normalizado

Con la aproximación seguida en este trabajo se consiguen resolver parte de los principales inconvenientes descritos en el uso de histogramas. Con este nuevo método basado en segmentación, las anomalías detectadas son dependientes del contexto. Sin embargo, como las medidas altas de actividad son en general poco frecuentes, estas siempre serán puntuadas con un valor alto de anomalía independientemente del segmento en el que se trabaje. Este hecho se puede visualizar comparando la Figura F.3.e con la Figura F.4.

F.1.4. DIAGRAMAS DE CAJAS

Esta técnica de detección de anomalías se basa en la construcción de un diagrama de cajas similar al representado en la Figura F.5 (Chen et al. 2013). Los diagramas de cajas son técnicas de estadística descriptiva basados en el cálculo de cuartiles y que permiten realizar la visualización de un conjunto de datos para extraer características de tendencia y/o simetría. Un diagrama de cajas se construye a través de las medidas extraídas de los propios datos calculando los cuartiles Q_1 (25%) y Q_3 (75%). La distancia entre los dos cuartiles es denominada distancia inter-cuartil (IQR) y es en base a la que se establecen los valores mínimos y máximos para la detección de anomalías. Una muestra que se encuentre una distancia $1,5 \cdot IQR$ por debajo de Q_1 o $1,5 \cdot IQR$ por encima de Q_2 será considerada como anómala. La región entre $Q_1 - 1,5 \cdot IQR$ y $Q_3 + 1,5 \cdot IQR$ contiene el 99,3% de las observaciones totales en una distribución de tipo normal. Así, sobre una distribución de tipo normal esta técnica es similar a la regla de 3σ descrita en la sección F.1.2.

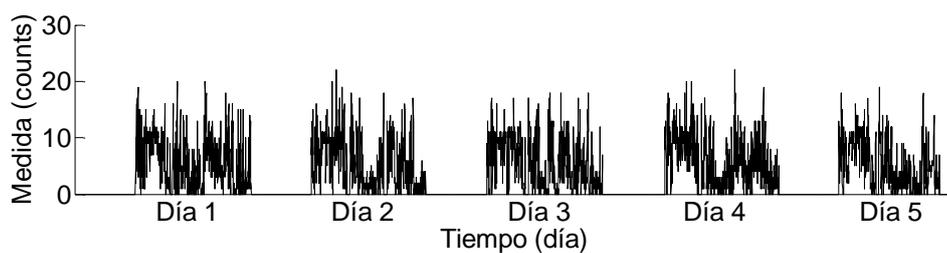


Figura F.5. Diagrama de cajas para la detección de anomalías

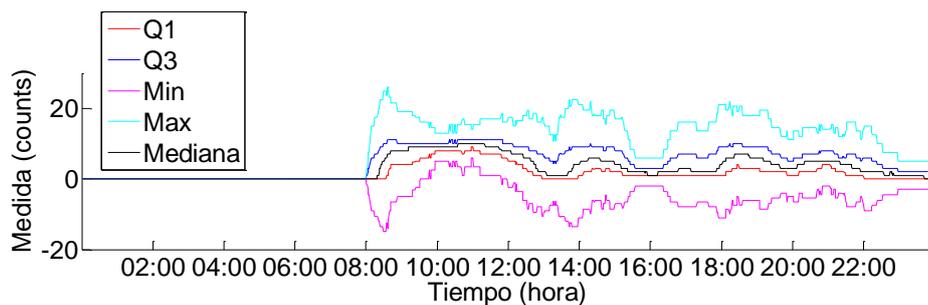
Partiendo de los mismos datos utilizados en el resto de secciones de este apéndice para la experimentación (ver Figura F.6.a), se construye la Figura F.6.b basada en el cálculo del diagrama de cajas para cada muestra de actividad física utilizando los segmentos de las medidas de los $N_r = 5$ días de actividad. Los parámetros del diagrama de cajas (mediana, Q1 y Q3) son calculados en la etapa de extracción de características (descrita en la sección 5.3.2.4) en el segmento correspondiente a un instante temporal n (segmento de $n - N_m/2$ a $n + N_m/2$) sobre los $N_r = 5$ días de actividad. En la Figura F.6.b se muestran las líneas que representan cada una de las características extraídas junto con el límite inferior y superior para la detección de anomalías. Utilizando el diagrama de cajas construido a partir de estos datos históricos, las anomalías son detectadas sobre las medidas del día de monitorización representado en la Figura F.3.d. La Figura F.6.b muestra la superposición de las líneas que definen el diagrama de cajas y la medida de la actividad en el día evaluable y que permite identificar visualmente los puntos detectados como anomalía en la Figura F.6.d mediante asteriscos negros.

El método propuesto basado en el uso de un diagrama de cajas con el que se modelaría el comportamiento normal, permite la detección de algunas anomalías pero no es capaz de detectar el comportamiento anómalo relacionado con el cambio en la hora de levantarse identificado por el usuario. Esto es debido a que en los instantes temporales en los que se produce el comportamiento anómalo, las medidas de actividad física tienen una gran desviación en el conjunto histórico utilizado. Esto hace que la distancia inter-cuartil sea elevada y que en consecuencia lo sea el margen de variación permitido para la detección de anomalías.

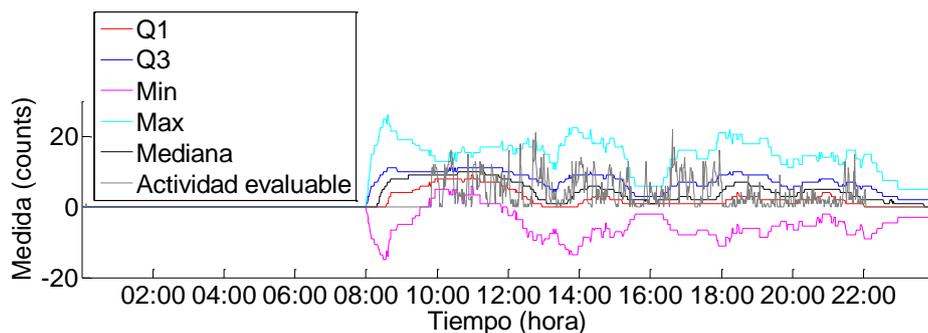
Aunque el uso de diagramas de cajas es una técnica ampliamente utilizada en el estado del arte tanto para detectar anomalías como para la valoración y análisis de procesos industriales, este método no es adecuado para el tipo de detección de anomalías deseadas en este trabajo. En el ejemplo de la Figura F.6 se muestra como un comportamiento identificado como anómalo no es detectado por el método basado en la técnica estudiada.



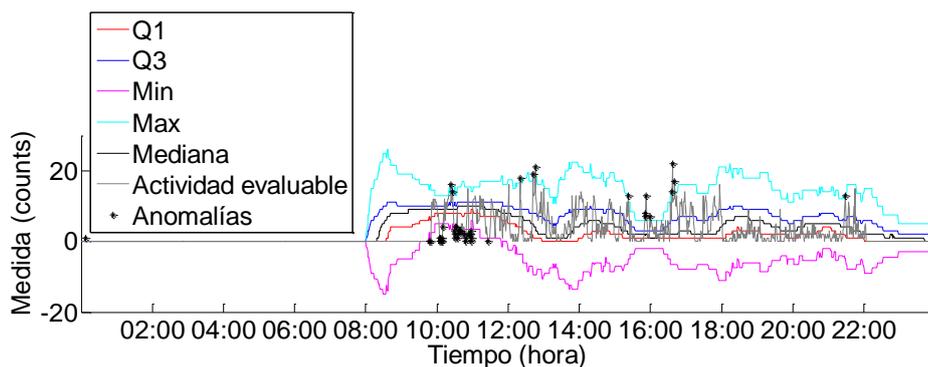
a) Medidas de actividad durante los cinco días de duración del experimento



b) Representación del diagrama de cajas



c) Representación del diagrama de cajas sobre un día evaluable



d) Detección de anomalías sobre el día evaluable

Figura F.6. Aplicación de diagramas de cajas para detección de anomalías

F.1.5. MEDIDAS DE SIMILITUD

El uso de una medida de similitud para la detección de anomalías parte de la existencia de un modelo esperado de comportamiento y el estudio de similitud que posee la medida de actividad del día evaluable con el modelo calculado. Dentro de las medidas de similitud destaca principalmente la distancia euclídea cuya expresión matemática para un espacio de dos dimensiones (x, y) se muestra en la Ecuación F.4. Este tipo de distancia representa la distancia “ordinaria” que se mediría a través de una regla entre dos puntos del espacio.

$$D(P_1, P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Ecuación F.4

Existen otros tipos de distancias derivados de la distancia euclídea con el objetivo de solventar algunas de sus principales limitaciones (altamente sensible a las unidades utilizadas y a variables correladas). Dentro de este conjunto de distancias puede mencionarse la distancia euclídea normalizada, la distancia de Mahalanobis (De Maesschalck, Jouan-Rimbaud, and Massart 2000) o la distancia de Chebyshev.

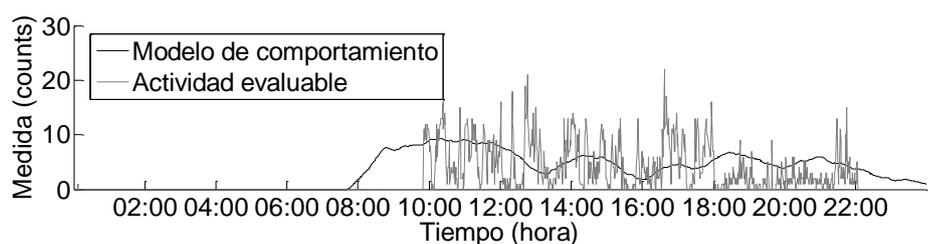
Otro tipo de medida de similitud es el cálculo de residuos. Los residuos de un modelo representan la parte de los datos estudiados que no pueden ser explicados mediante el modelo establecido (diferencia entre el modelo y las observaciones). La Ecuación F.5 muestra la expresión matemática del cálculo de los residuos (ε) donde y representa el valor de la medida y \hat{y}_i el valor del modelo para la muestra i .

$$\varepsilon_i = y_i - \hat{y}_i$$

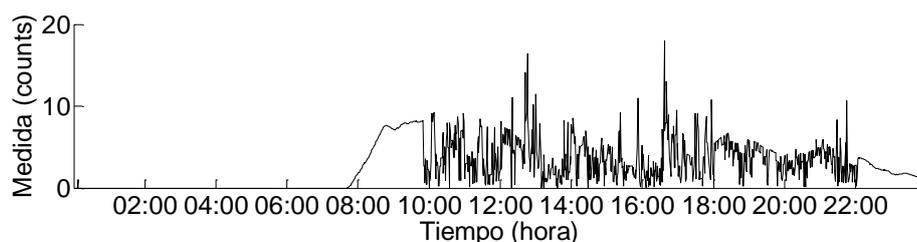
Ecuación F.5

Adaptando la expresión de la distancia euclídea (Ecuación F.4) para un problema unidimensional como el que se plantea en este trabajo, se puede apreciar como el cálculo de la distancia euclídea es similar al cálculo del valor absoluto de los residuos (Ecuación F.5).

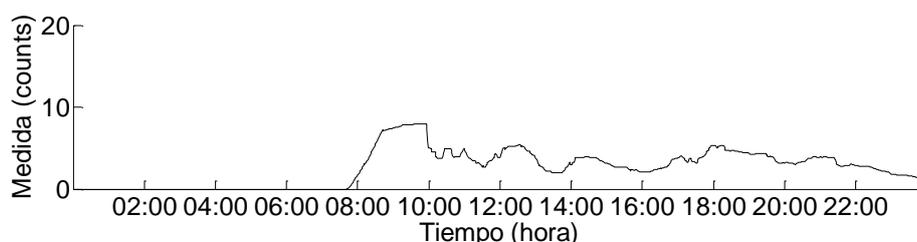
Siguiendo con el mismo ejemplo utilizado en los apartados anteriores, la Figura F.7 muestra el uso de la distancia euclídea (valor absoluto del cálculo de los residuos) para calcular para cada muestra de la medida de actividad en el día a estudiar su similitud con el modelo esperado. Ambas variables (medida y modelo) son representadas en la Figura F.7.a y el valor absoluto de los residuos calculados en la Figura F.7.b.



a) Representación del modelo y de la medida de actividad del día evaluable



b) Distancia euclídea punto a punto



c) Distancia euclídea punto a punto filtrada

Figura F.7. Aplicación de la distancia euclídea como medida de similitud

Como ya se ha mencionado en las secciones anteriores, la medida de actividad del usuario difiere del modelo principalmente en el comienzo de su actividad (alrededor de las 10:00 en lugar de las 08:00) y en un descenso de actividad durante la tarde (actividad alta de 18:30 a 20:30 de acuerdo a su rutina y reposo en el día evaluable).

Los propios valores obtenidos en el cálculo de la distancia euclídea podrían ser utilizados para proporcionar una puntuación de anomalía para cada muestra temporal estudiada. Con el objetivo de obtener una medida más clara, la Figura F.7.c muestra la aplicación de un filtrado de mediana sobre la distancia calculada. Este filtro permite eliminar el comportamiento impulsivo en la variable y visualizar de forma más clara la tendencia que sigue la diferencia de actividad.

En general, el principal inconveniente del uso de una medida de distancia para la detección de anomalías es el establecimiento de umbrales para a partir de esa distancia permitan determinar qué muestras se corresponden o no con comportamientos no esperados. La magnitud de la distancia euclídea no es por sí misma auto-explicativa ya

que una distancia alta tendrá más o menos impacto dependiendo del orden de las variables implicadas.

Así, la distancia euclídea debido a su relación con el cálculo de los residuos es una buena medida de similitud entre una medida y un modelo. Sin embargo, para su uso en la detección de anomalías de comportamiento es necesaria su combinación con otro tipo de método que permita identificar las distancias significantes. A su vez, al trabajar en valor absoluto, la distancia euclídea no permite conocer el sentido de la anomalía (decremento o incremento anómalo de actividad) y hace que el cálculo de los residuos suponga una medida mucho más interesante.

F.2. ANÁLISIS DE MÉTODOS DE EVALUACIÓN DE ANOMALÍAS

F.2.1. INTRODUCCIÓN

En esta sección se analiza el uso de tres métodos para la evaluación de anomalías: uso de la desviación estándar del modelo, uso de la desviación estándar de la medida y funciones de pertenencia Fuzzy. Este análisis se realiza después de haber seleccionado el cálculo de los residuos como medida de similitud en la fase I del método propuesto (ver sección 6.3.3.2) y el objetivo es identificar el método de evaluación más adecuado para la detección de anomalías. El método seleccionado es implementado en la fase III del método (ver sección 6.3.3.4).

Para realizar la detección y evaluación de anomalías, una vez que se selecciona el cálculo de los residuos como medida de similitud, es necesario fijar un umbral. Una de las aproximaciones más sencillas es el uso de umbrales estáticos basados en conocimiento experto o en una etapa previa de entrenamiento. En estos casos, un cambio en las medidas o en la unidad de las mismas, implicaría la necesidad de realizar un reajuste para la detección. Una aproximación dinámica de establecimiento de umbrales evita este tipo de inconvenientes y permite la selección de forma individual y adaptativa para cada usuario e instante temporal.

F.2.2. USO DE LA DESVIACIÓN ESTÁNDAR DEL MODELO

La primera de las aproximaciones utilizadas de establecimiento de umbrales para la detección de anomalías fue el uso de la desviación estándar del modelo esperado. La Figura F.8 muestra el establecimiento de un umbral basado en la desviación estándar del modelo esperado. Este umbral se representa sobre el valor absoluto de la salida de la fase I del método, equivalente a la distancia euclídea filtrada entre la medida de actividad y el modelo.

La desviación estándar de una distribución de datos x de media \bar{x} se calcula siguiendo la expresión de la Ecuación F.6. Se trata de una medida de dispersión que indica como los datos se distribuyen respecto a la media y puede ser utilizada como grado de incertidumbre para analizar la posibilidad de que una determinada muestra pertenezca a la distribución estudiada.

$$\sigma = \sqrt{\frac{1}{N-1} \cdot \sum_{i=1}^N (x_i - \bar{x})^2}$$

Ecuación F.6

La técnica utilizada en la Figura F.8 permitiría detectar las anomalías esperadas de acuerdo al comportamiento del usuario en el día de estudio respecto al modelo. Los comportamientos anómalos (tramos temporales de comportamiento anómalo) podrían ser identificados mediante un algoritmo que una las anomalías puntuales que tienen lugar de forma secuencial. Sin embargo, el método propuesto no detectaría anomalías de forma contextual y adaptativa, ya que se posee el mismo umbral de detección para todos los instantes temporales. La detección tiene lugar utilizando el mismo valor umbral en los instantes en los que se espera un comportamiento muy estable que en aquellos en los que se espera un comportamiento más variable.

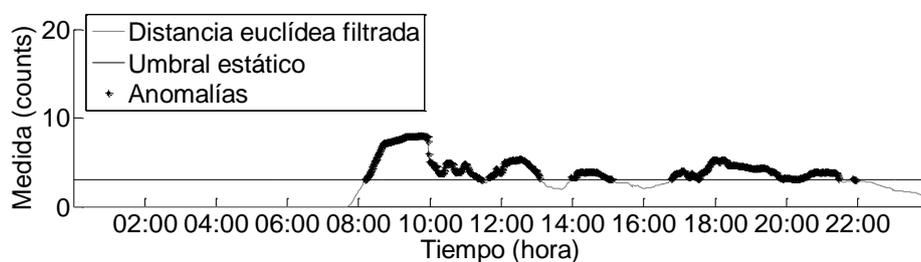


Figura F.8. Detección mediante umbral estático

Para resolver esta limitación, se propone el uso de un umbral adaptativo calculado a través de la etapa de segmentación. Para cada instante n , el umbral de detección es calculado mediante los segmentos temporales extraídos en los $N_r = 5$ días anteriores. La Figura F.9 muestra la detección de anomalías siguiendo este procedimiento. Puesto que el modelo de comportamiento es construido a partir de la media móvil de la actividad en los $N_r = 5$ días anteriores, la desviación estándar usada como umbral simboliza como las muestras medidas se adaptan al valor obtenido en el modelo.

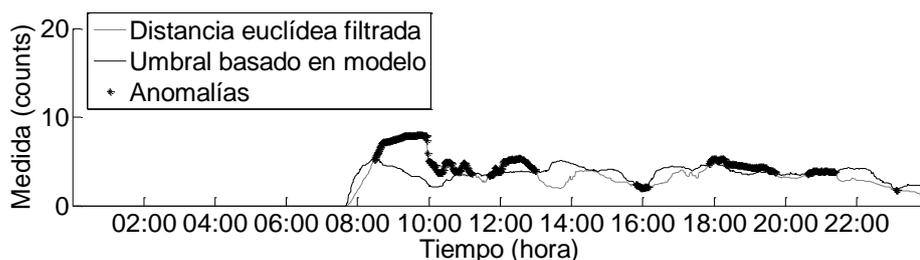


Figura F.9. Detección mediante umbral basado en la desviación del modelo

El uso de este nuevo umbral para la detección permite identificar anomalías contextuales como las muestras en torno a las 21:00 que no habían sido identificadas como anomalías anteriormente. Sin embargo, este método sigue teniendo limitaciones ya que el usuario se acuesta con anterioridad al resto de días (22:00 en lugar de 23:00) y no se detecta ninguna muestra anómala en ese tramo temporal. Igual que ocurría en la sección F.1.4 en el cálculo de los diagramas de cajas, la desviación existente en las medidas durante ese tramo temporal es tan alta que hace que un gran número de valores de medida sean considerados como normales. Para que una anomalía fuera detectada en ese tramo temporal mediante la técnica propuesta debería poseer una diferencia muy alta.

F.2.3. USO DE LA DESVIACIÓN ESTÁNDAR DE LA MEDIDA

Otra posible estrategia para el establecimiento de los umbrales de detección es el uso de la desviación de la medida en lugar de la desviación del modelo esperado. De la misma forma que en la Figura F.9, la Figura F.10 muestra la detección cuando el umbral se establece como la desviación calculada en cada segmento temporal (de $n - N_m/2$ a $n + N_m/2$ para cada muestra n) de la medida de actividad. En este caso, no solo se detecta una anomalía de mayor duración durante la actividad matutina del usuario, sino que el método es capaz de identificar el cambio de actividad nocturna.

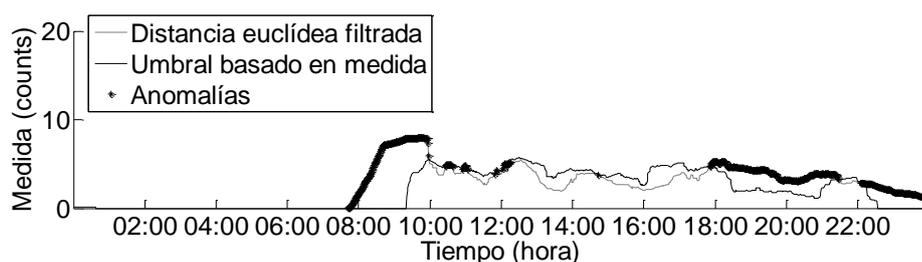


Figura F.10. Detección mediante umbral basado en la desviación de la medida

Del estudio de estos dos últimos métodos para el establecimiento de umbrales de detección (uso de la desviación estándar del modelo y uso de la desviación estándar de la medida) se aprecia como un compromiso entre ambos permitiría obtener mejores resultados que su uso individual. Para un determinado instante de medida n si la desviación estándar del modelo es alta, un gran rango de medidas de actividad serán consideradas como normales. Sin embargo, si esa medida de actividad se corresponde con una desviación muy baja, la muestra debería ser etiquetada como anómala. Conseguir el compromiso entre estos dos umbrales, será uno de los puntos clave en el uso de los residuos para la detección de anomalías.

F.2.4. FUNCIONES DE PERTENENCIA FUZZY

Una de las principales ventajas del uso de los residuos del modelo para la detección de anomalías es que la propia magnitud calculada puede ser utilizada como puntuación de anomalía. Sin embargo, puesto que el rango de valores que se pueden obtener no está acotado, la detección y comparación entre anomalías es poco intuitiva y difícilmente interpretable.

Una posible solución es el uso de técnicas de lógica difusa o lógica Fuzzy. Estas técnicas se basan en la creación de un conjunto difuso A con valores comprendidos entre 0 y 1 y una función de pertenencia $\mu_A(x)$ que permita el mapeo entre el conjunto de observaciones (X) y el conjunto difuso tal y como muestra la Ecuación F.7.

$$\mu_A(x): X \rightarrow [0,1]$$

Ecuación F.7

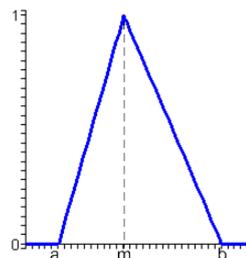
Aplicando esta técnica para la detección de anomalías, se debe establecer una función de pertenencia Fuzzy que dada la distancia genere un valor entre 0 y 1 identificando el grado de anomalía. Un valor próximo a 1 identificará una muestra anómala y un valor próximo a cero una muestra normal.

Existen diferentes tipos de funciones de pertenencia que pueden ser utilizadas para el mapeo de las observaciones en el conjunto difuso. A continuación, con imágenes extraídas de (Olmos Castillo 2008), se muestran las principales alternativas:

- *Función triangular*. Se define mediante el límite inferior (a), el límite superior (b) y el valor máximo (m). Su expresión matemática y su representación gráfica se muestran en la Figura F.11.

$$\mu_A(x) = \begin{cases} 0 & \text{Si } x \leq a \\ \frac{x-a}{m-a} & \text{Si } a < x \leq m \\ \frac{b-x}{b-m} & \text{Si } m < x < b \\ 0 & \text{Si } x \geq b \end{cases}$$

a) Expresión matemática



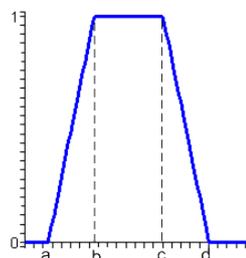
b) Representación gráfica

Figura F.11. Función de pertenencia triangular

- *Función trapezoidal*. Se define mediante cuatro parámetros a , b , c y d que deben cumplir la condición: $a \leq b \leq c \leq d$. Existen versiones características de esta función en las que algunos de los parámetros toman valores no finitos. Su expresión matemática y su representación gráfica se muestran en la Figura F.12.

$$\mu_A(x) = \begin{cases} 0 & \text{Si } x \leq a \\ \frac{x-a}{b-a} & \text{Si } a \leq x \leq b \\ 1 & \text{Si } b \leq x \leq c \\ \frac{d-x}{d-c} & \text{Si } c \leq x \leq d \end{cases}$$

a) Expresión matemática

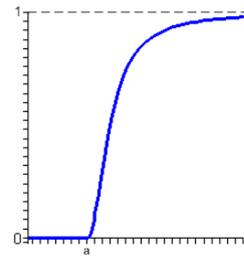


b) Representación gráfica

Figura F.12. Función de pertenencia trapezoidal

- *Función gamma*. Se define mediante un límite inferior (a) y una constante k ($k > 1$). Esta función de caracteriza por un rápido crecimiento a partir del valor a que será más rápido cuanto mayor sea el valor del parámetro k . Esta función de pertenencia presenta una asíntota en el valor 1, por lo que nunca se tomará este valor en el conjunto difuso. Su expresión matemática y su representación gráfica se muestran en la Figura F.13.

$$\mu_A(x) = \begin{cases} 0 & \text{Si } x \leq a \\ \frac{k \cdot (x - a)^2}{a + k \cdot (x - a)^2} & \text{Si } x > a \end{cases}$$



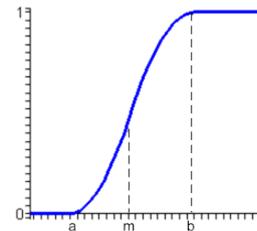
a) Expresión matemática

b) Representación gráfica

Figura F.13. Función de pertenencia gamma

- *Función sigmoideal.* Se define a través de tres parámetros a (límite inferior), b (límite superior) y m (punto de inflexión) que deben cumplir la condición: $a \leq m \leq b$. El crecimiento es más lento cuando mayor será la distancia entre los parámetros a y b . Su expresión matemática y su representación gráfica se muestran en la Figura F.14.

$$\mu_A(x) = \begin{cases} 0 & \text{Si } x \leq a \\ 2 \cdot \left[\frac{x - a}{b - a} \right]^2 & \text{Si } a < x \leq m \\ 1 - 2 \cdot \left[\frac{x - b}{b - a} \right]^2 & \text{Si } m < x < b \\ 1 & \text{Si } x \geq b \end{cases}$$



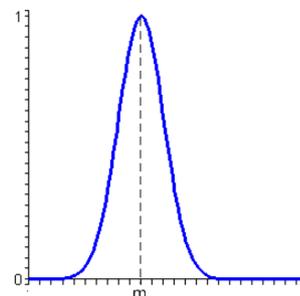
a) Expresión matemática

b) Representación gráfica

Figura F.14. Función de pertenencia sigmoideal

- *Función gaussiana.* Se define mediante la media (m) y el parámetro k relacionado con la anchura que toma la campana. Cuanto mayor sea el valor de k , más estrecha será la función de pertenencia. Su expresión matemática y su representación gráfica se muestran en la Figura F.15.

$$\mu_A(x) = e^{-k \cdot (x - m)^2}$$



a) Expresión matemática

b) Representación gráfica

Figura F.15. Función de pertenencia gaussiana

La selección de una u otra función de pertenencia dependerá de la relación deseada entre la distancia calculada y el grado de anomalía. Una vez seleccionada la forma de la función de pertenencia es necesario establecer los parámetros a partir de los cuales se construya la función.

F.3. LISTA DE USUARIOS

Este apéndice recoge la información de los 12 usuarios involucrados en el proceso de validación del método de detección de comportamientos anómalos presentado en la sección 6.3.3 del Capítulo 6. Los resultados obtenidos son descritos y analizados en la sección 6.3.4.

La Tabla F.1 muestra para cada usuario involucrado en la experimentación su sexo, edad, peso, estatura e índice de masa corporal (BMI). Cada uno de los usuarios fue monitorizado durante dos días cuya separación temporal y duración del experimento se muestra en esta misma tabla. El perfil de actividad mostrado está relacionado con el tipo de ejercicio físico realizado durante el periodo de monitorización. Los usuarios con perfil de actividad bajo no realizaron ningún tipo de ejercicio físico y los usuarios con perfil de actividad alto llevaron a cabo actividades deportivas de alta intensidad.

Tabla F.1. Lista de usuarios involucrados en la validación del método de detección

Nº	Sexo	Edad	Peso (Kg)	Estatura (m)	BMI	Experimento		Perfil de actividad
						Duración (min)	Separación (días)	
1	M	28	57,0	1,70	19,72	1560	21	Medio
2	M	29	76,0	1,63	28,60	1560	21	Bajo
3	M	26	45,0	1,56	18,49	1560	18	Bajo
4	M	28	69,0	1,69	24,16	1740	20	Bajo
5	M	29	61,0	1,76	19,69	1800	19	Medio
6	V	33	73,5	1,77	23,46	1380	15	Alto
7	V	34	66,0	1,70	22,84	1380	15	Medio
8	M	29	57,0	1,69	19,96	1380	22	Alto
9	V	28	81,0	1,78	25,56	1560	17	Alto
10	V	27	61,0	1,75	19,92	1680	39	Alto
11	V	30	83,0	1,86	23,99	1860	32	Alto
12	V	30	80,0	1,75	26,12	1200	22	Bajo

La Tabla F.2 muestra los comportamientos identificados como anómalos por los usuarios en base a los cuestionarios de actividad cubiertos. En esta tabla se muestra el nivel de actividad reportado en los cuestionarios para el primer día (nivel de actividad día anterior) y para el segundo día de monitorización (nivel de actividad día evaluable). Junto con cada comportamiento se muestra la explicación dada por el usuario para justificar su elección.

Tabla F.2. Comportamientos anómalos identificados por los usuarios

Usuario	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día evaluable	Nivel actividad día anterior
1	20:00	21:00	+1	1	0
	<i>El primer día quitó el reloj para limpiar a esa hora</i>				
1	23:00	00:00	-1	2	1
	<i>El segundo día de monitorización salió a cenar fuera y estuvo tomando algo con amigos</i>				
5	20:00	21:00	-1	1	2
	<i>El segundo día de monitorización no fue al gimnasio</i>				
8	14:00	15:00	-1	1	3
	<i>A la hora de la comida, el primer día que fue monitorizado practicó actividades deportivas y el segundo día estuvo en reposo en ese tramo horario</i>				
8	19:00	21:00	+1	3	1
	<i>Realizó actividades deportivas en ese tramo horario durante el segundo día de monitorización</i>				
8	22:00	02:00	-1	0	1
	<i>Se acostó antes que el primer día que fue monitorizado</i>				
9	18:00	19:00	-1	1	2
	<i>El primer día de monitorización anduvo en bicicleta durante este tramo y el segundo día estuvo en reposo viendo la televisión</i>				
9	20:00	00:00	+1	2	0
	<i>El primer día quitó el dispositivo de monitorización durante este tramo horario y el segundo día lo dejó puesto mientras cocinaba y hacía labores del hogar</i>				
12	19:00	21:00	-1	0	1
	<i>Durmió la siesta en este tramo horario durante el segundo día de monitorización</i>				

Analizando el parámetro V del método de detección, la Tabla F.3 muestra los comportamientos detectados correctamente (TP), los comportamientos detectados no reales (FP) y los comportamientos ocurridos que el método no es capaz de detectar (FN). Variando el parámetro V se consigue determinar que con $V = 1$ la precisión y sensibilidad del método es máxima.

Tabla F.3. Resultados del método de detección variando el parámetro V

Usuario	V = 0			V = 0,25			V = 0,50			V = 0,75			V = 1		
	TP	FP	FN	TP	FP	FN	TP	FP	FN	TP	FP	FN	TP	FP	FN
1	2	1	0	2	1	0	2	1	0	2	1	0	2	0	0
2	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
3	0	4	0	0	3	0	0	1	0	0	1	0	0	0	0
4	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
5	1	2	0	1	2	0	1	1	0	1	0	0	1	0	0
6	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	3	1	0	3	1	0	3	1	0	3	0	0	3	0	0
9	2	1	0	2	0	0	2	0	0	2	0	0	2	0	0
10	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
11	0	2	0	0	1	0	0	1	0	0	1	0	0	0	0
12	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
TOTAL	9	15	0	9	11	0	9	5	0	9	3	0	9	0	0

APÉNDICE G.

PROCEDIMIENTO DE EXPERIMENTACIÓN

G.1. INTRODUCCIÓN

Este apéndice describe el procedimiento utilizado para la experimentación durante las validaciones llevadas a cabo en este trabajo. Este procedimiento ha sido seguido para las validaciones descritas en los Capítulo 4, Capítulo 5, Capítulo 6 y Capítulo 7. En las siguientes secciones se describe el cuestionario de actividad que se solicita a cada usuario que cubra para conocer su perfil de actividad y el consentimiento informado que se les facilitó para informarles sobre su participación en este trabajo.

G.2. CUESTIONARIO DE ACTIVIDAD

A los usuarios participantes en la validación se les pidió que llevaran puesto el dispositivo vestible para la monitorización durante las 24 horas del día siguiendo su rutina habitual de actividad. Se les permitió quitar el dispositivo cuando se asearan y cuando durmieran. En el caso de usuarios que suelen dormir con reloj, se les permitió que no lo quitasen. El objetivo es que sigan su rutina diaria sin realizar modificaciones de la misma por el hecho de estar siendo monitorizados. La monitorización se realizó de forma transparente para el usuario descargando los datos automáticamente a un servidor central para su procesamiento (modelado) y análisis (detección).

Para cada día de monitorización, se les entregó una tabla similar a la Tabla G.1 utilizada en (Craig et al. 2003) entre otros. Para cada una de las horas del día, se le pide al usuario que indique el nivel de actividad realizado: nulo, bajo, medio, alto o muy alto.

- *Nulo (valor 0)*. Se corresponde con periodos de tiempo en los que el usuario no portó el reloj o que se encontró durmiendo.
- *Bajo (valor 1)*. Periodos de tiempo en que el usuario está en reposo.
- *Medio (valor 2)*. Se corresponde con actividad que requieren poco esfuerzo como caminar o realizar las tareas del hogar.
- *Alto (valor 3)*. Actividades deportivas de intensidad media: caminar, correr, escalar, andar en bicicleta...
- *Muy alto (valor 4)*. Actividades deportivas de alta intensidad: atletismo, entrenamientos guiados...

Tabla G.1. Cuestionario de actividad física

Hora	Actividad nula	Actividad baja	Actividad media	Actividad alta	Actividad muy alta
00:00					
01:00					
02:00					
03:00					
...					
...					
...					
22:00					
23:00					

En base a los valores indicados en los cuestionarios, para cada hora puede construirse una variable similar a la mostrada en la Figura G.1 que representa el comportamiento reportado en la Tabla G.2. Esta variable identifica la actividad teórica realizada según el cuestionario cubierto y se construye identificando con un número (0, 1, 2, 3 y 4) el nivel de actividad reportado por el usuario.

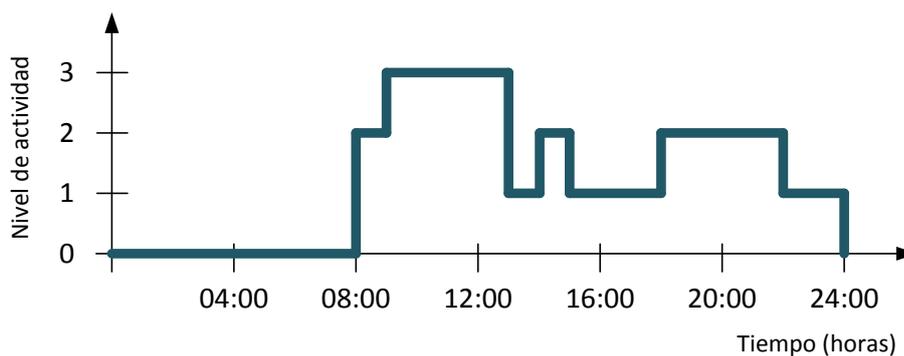


Figura G.1. Ejemplo de construcción de actividad teórica

Tabla G.2. Ejemplo de cuestionario de actividad física

Hora	Actividad nula	Actividad baja	Actividad media	Actividad alta	Actividad muy alta
00:00	X				
01:00	X				
02:00	X				
03:00	X				
04:00	X				
05:00	X				
06:00	X				
07:00	X				
08:00			X		
09:00				X	
10:00				X	
11:00				X	
12:00				X	
13:00		X			
14:00			X		
15:00		X			
16:00		X			
17:00		X			
18:00		X			
19:00			X		
20:00			X		
21:00			X		
22:00		X			
23:00		X			

G.3. CONSENTIMIENTO INFORMADO

A cada uno de los usuarios participantes en la validación se le entregó el siguiente documento informándoles sobre el objetivo del proyecto en el que colaboraban y los procedimientos que se seguirían en el tratamiento de la información personal recogida durante la experimentación.

INFORMACIÓN DIRIGIDA A LOS PARTICIPANTES POTENCIALES DEL PROYECTO:

“MÉTODO AUTOMÁTICO Y ADAPTATIVO PARA LA DETECCIÓN DE ANOMALÍAS EN LA ACTIVIDAD FÍSICA MEDIANTE UN SENSOR DE ACELERACIÓN NO INVASIVO”

El objetivo del proyecto es la monitorización de la actividad física en personas mayores para la detección a partir de su medida de posibles comportamientos anómalos. Por comportamiento anómalo se entiende la detección de comportamientos no esperados de acuerdo a la rutina de la persona monitorizada. La detección de estos comportamientos permitirá en el futuro la detección, seguimiento y control de posibles trastornos relacionados con la actividad.

En la fase del proyecto en la que nos encontramos, buscamos voluntarios con el objetivo de monitorizar su actividad física para a partir de su procesamiento y análisis trabajar en el diseño de un método automático y adaptativo que permita la detección de anomalías. La monitorización de la actividad física tendrá lugar mediante un sensor de aceleración integrado en un reloj con apariencia deportiva.

Durante el tiempo que dure la experimentación, deberá vestir el reloj durante las 24 horas del día. Puede quitarlo para dormir, asearse y en general para cualquier actividad en la que interfiera negativamente. Debe seguir su rutina diaria sin realizar modificaciones por el hecho de estar siendo monitorizado. En determinados casos, puede que se facilite algún tipo de equipamiento informático adicional que se colocará en su hogar.

Su representante y persona de contacto es: XXXXXXXX

CONSENTIMIENTO INFORMADO DE PARTICIPACIÓN EN EL PROYECTO

El proyecto “Método automático y adaptativo para la detección de anomalías en la actividad física mediante un sensor de aceleración no invasivo” necesita que proporcione determinada información sobre su persona para el proyecto. Nos gustaría poseer información personal sobre usted para conocer su rutina de comportamiento y comprender cómo usa, percibe, o se ve afectado por el uso de las nuevas tecnologías.

Para ello, usted decide la información que desea facilitarnos. Su nombre y dirección no se incluirán en los resultados de las pruebas o en las entrevistas que se lleven a cabo. A su vez, la información monitorizada y recogida por el sensor de aceleración utilizando en el proyecto será almacenada de forma anónima en una base de datos segura albergada por Fundación CTIC – Centro tecnológico. Toda la información se codificará y se mantendrá completamente confidencial; es decir, no será posible identificarle como persona sin un código que sólo conocerá su representante local.

Toda la información recogida y almacenada será procesada y analizada con el objetivo de investigar y desarrollar métodos que permitan la detección de comportamientos anómalos en personas mayores.

Me han explicado el proyecto “Método automático y adaptativo para la detección de anomalías en la actividad física mediante un sensor de aceleración no invasivo” y he tenido la oportunidad de hacer preguntas sobre él. También he sido informado de que puedo hacer preguntas sobre el proyecto y sobre mi propia participación en el mismo en cualquier momento que desee. He leído la información relativa al proyecto que el representante me ha proporcionado y la entiendo.

Entiendo que mi participación en este proyecto es totalmente voluntaria y que puedo cesar mi participación en cualquier momento que desee sin tener que proporcionar ninguna explicación y sin incurrir en ningún tipo de consecuencia negativa.

También entiendo que toda la información personal recogida durante el proyecto se mantendrá confidencial para siempre. Estoy de acuerdo en que se utilice la información personal sobre mí con fines de investigación y con los objetivos mencionados anteriormente.

Teniendo en cuenta lo anterior, doy mi consentimiento para participar en el proyecto de investigación “Método automático y adaptativo para la detección de anomalías en la actividad física mediante un sensor de aceleración no invasivo”.

Fecha, nombre y firma:

APÉNDICE H.

USUARIO DE EJEMPLO:

JOSÉ GARCÍA

H.1. INTRODUCCIÓN

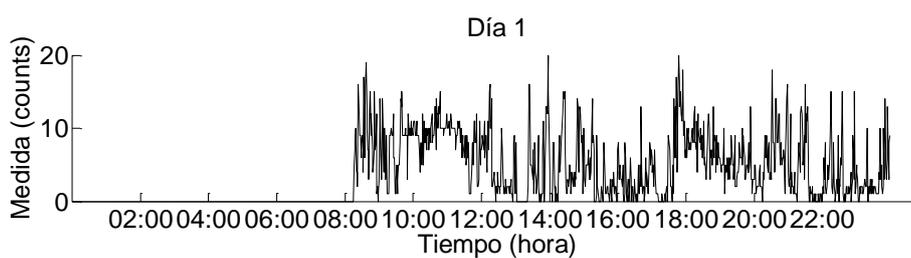
Este apéndice describe a un usuario de ejemplo al que se ha decidido llamar José García y cuya actividad física ha sido utilizada durante este trabajo para el estudio y análisis de algunos métodos de trabajo. En las siguientes secciones se describe su rutina y se muestra el funcionamiento de los métodos propuestos.

H.2. DESCRIPCIÓN DE LA RUTINA DEL USUARIO

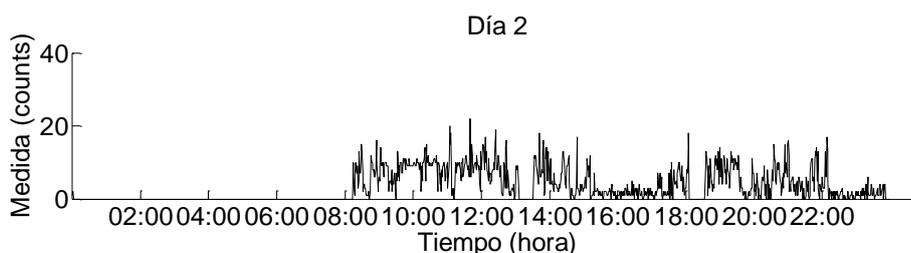
José García se levanta habitualmente alrededor de las 8 de la mañana y tras desayunar y hacer alguna tarea del hogar, sale a caminar a ritmo fuerte durante al menos 1 hora. A continuación suele hacer ejercicios al aire libre utilizando los aparatos gimnásticos de los que se dispone en los “Parques para mayores” de su ciudad. Después de hacer deporte vuelve a casa sobre las 1 de la tarde para descansar un poco y ayudar a su mujer en las tareas del hogar. Come todos los días en torno a las 3 de la tarde y a continuación navega por Internet o duerme la siesta hasta las 6 de la tarde que vuelve de nuevo a salir para encontrarse con amigos y dar un paseo por la ciudad con su mujer. Cena a las 10 de la noche y suele acostarse a las 12 tras ver un rato la televisión.

H.3. EJEMPLO DE APLICACIÓN DEL MÉTODO DE MODELADO PROPUESTO

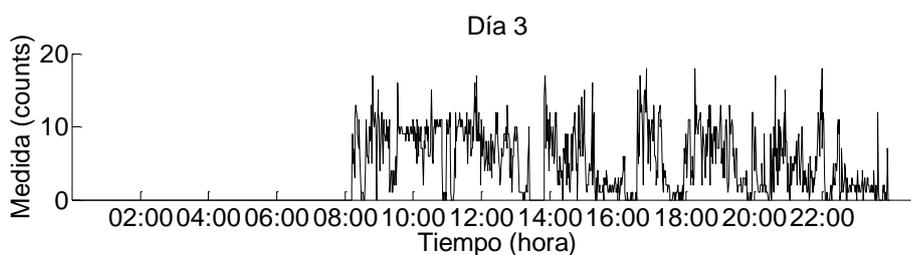
La Figura H.1 muestra la actividad física para un usuario de ejemplo durante los $N_r = 5$ días que serán utilizados para la aplicación del método de modelado propuesto (ver sección 5.3). Para ello, sobre cada una de las medidas representadas en la Figura H.1 se aplica el método de acondicionamiento descrito en la sección 5.3.2. En primer lugar se ejecuta la etapa de pre-procesamiento para la detección y tratamiento de medidas redundantes y ausentes. Para cada uno de los días de medida se obtiene una variable de 1440 muestras que posteriormente es tratada en la etapa de segmentación y extracción de características. El resultado, para cada variable medida, se muestra en la Figura H.2. Se aprecia como la salida de la etapa de acondicionamiento es similar a la aplicación de un filtro de media sobre la medida original de actividad, tal y como se describió en 5.3.2.5.



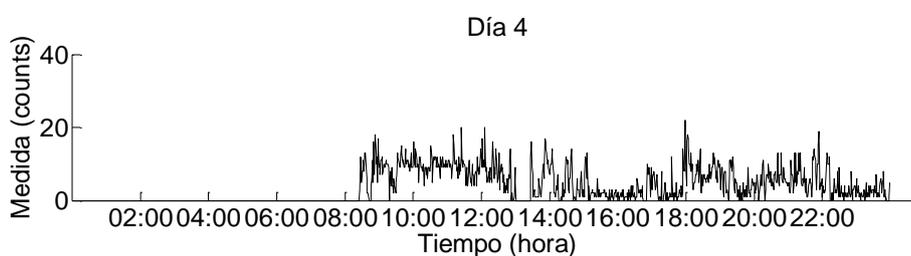
a) Medida de actividad física del Día 1



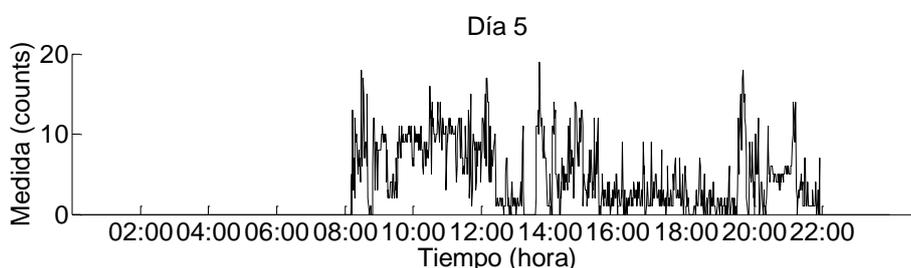
b) Medida de actividad física del Día 2



c) Medida de actividad física del Día 3

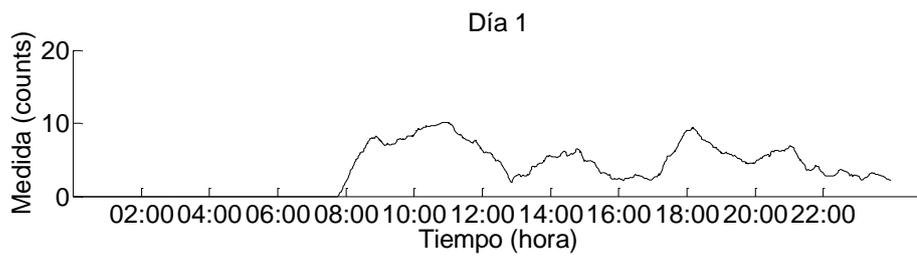


d) Medida de actividad física del Día 4



e) Medida de actividad física del Día 5

Figura H.1. Medidas de actividad física durante $N_r = 5$ días



a) Medida de actividad física acondicionada del Día 1



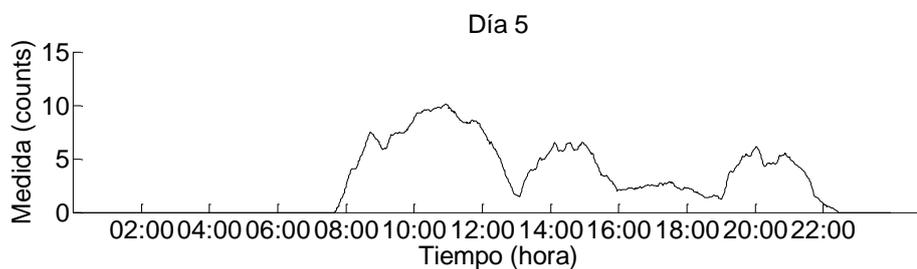
b) Medida de actividad física acondicionada del Día 2



c) Medida de actividad física acondicionada del Día 3



d) Medida de actividad física acondicionada del Día 4



e) Medida de actividad física acondicionada del Día 5

Figura H.2. Medidas de actividad física acondicionadas durante $N_r = 5$ días

A partir de la medida de la actividad física acondicionada sobre los $N_r = 5$ últimos días y mediante el método descrito en la sección 5.3.3, se identifica la rutina (modelo esperado) representada en la Figura H.3. Esta variable permite representar la rutina del usuario durante los 5 días utilizados para su cálculo y que según varios trabajos como (Heil et al. 2012) (Trost et al. 2005) y (Fouquet and Franco 2010) sirve para definir la rutina a nivel general del usuario monitorizado. En el caso del ejemplo del modelo obtenido en la Figura H.3, tal y como se describe en la sección H.1, representa adecuadamente la rutina del usuario.

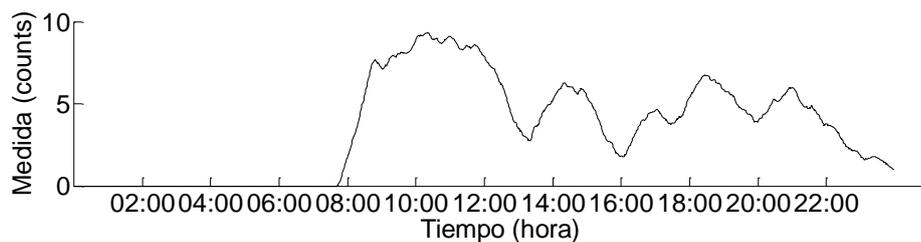


Figura H.3. Salida del método de modelado: rutina

Teniendo en cuenta la rutina del usuario y el modelo de comportamiento obtenido, puede apreciarse el nivel de actividad alto de 09:00 a 12:00 correspondiente con el ejercicio realizado por el usuario, la bajada del nivel a las 13:00, la realización de tareas del hogar de 14:00 a 15:00, el descenso de actividad cuando come y descansa de 15:00 a 17:00 y el aumento de 18:00 a 22:00 cuando pasea y se encuentra con amigos. El valle de actividad presente en el modelo a las 20:00 hace presagiar que es a esa hora cuando suele quedar con ellos tras dar un paseo por la ciudad.

H.4. EJEMPLO DE APLICACIÓN DEL MÉTODO DE DETECCIÓN PROPUESTO

Partiendo de la rutina identificada en la sección H.3, esta sección muestra el ejemplo de aplicación del método propuesto para la detección de anomalías (ver sección 6.3) sobre el Día 6 de monitorización. La actividad física medida durante ese día se representa en la Figura H.4. La medida de actividad utilizada es la perteneciente a un día de fin de semana en el que el usuario se levanta más tarde de lo habitual (decremento de actividad anómalo detectado entre las 08:00 y las 10:00 de la mañana y en el que modifica su actividad nocturna. Los fines de semana el usuario duerme la siesta y ve la televisión entre las 18:00 y las 21:00 y se acuesta más temprano de lo habitual (decremento de actividad detectado a las 22:00). Estos comportamientos anómalos que el método propuesto debería ser capaz de detectar se detallan en la Tabla H.1.

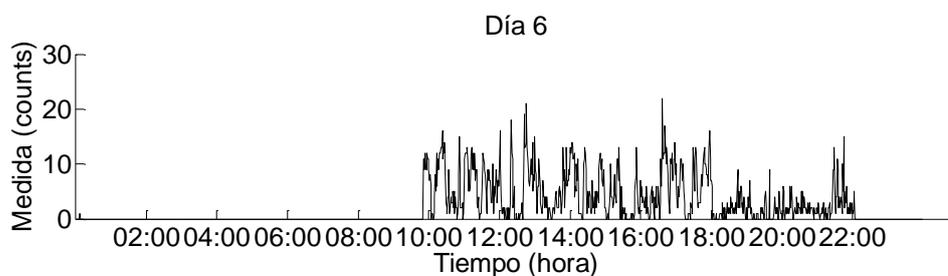


Figura H.4. Medida de actividad física del Día 6 de monitorización

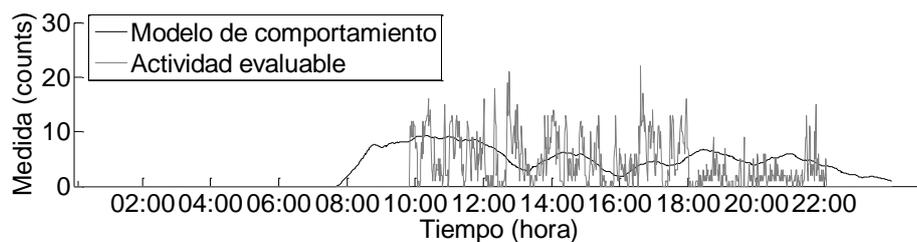
Tabla H.1. Comportamientos anómalos identificados en el Día 6

Comportamiento anómalo	Tiempo inicial	Tiempo final	Descripción
1	08:00	10:00	<i>El usuario se despierta más tarde que lo habitual</i>
2	18:30	20:30	<i>El usuario está en reposo viendo la televisión y durmiendo la siesta</i>
3	21:30	00:00	<i>El usuario se acuesta antes de lo que acostumbra</i>

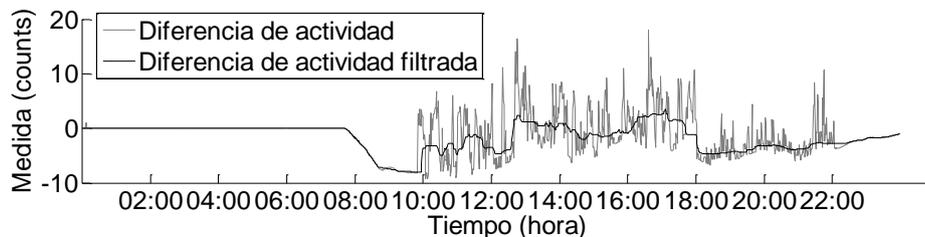
El ejemplo de aplicación del método propuesto para la detección de anomalías (ver sección 6.3) se muestra en la Figura H.5. El modelo de comportamiento calculado (línea negra) se muestra en la Figura H.5.a junto con la medida de actividad del día en el que se desean detectar anomalías (línea gris). La diferencia de actividad es calculada en la fase I del método propuesto a partir del cálculo de los residuos del modelo (diferencia de actividad). El resultado se muestra en la Figura H.5.b (línea gris) junto con la salida

de esta fase (línea negra) tras la aplicación de un proceso de filtrado de mediana de $N_m = 60$ muestras.

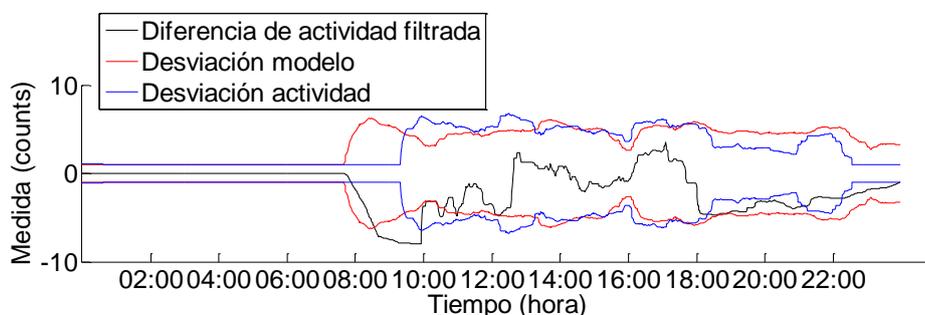
El resultado de la fase II (extracción de características estadísticas) se muestra en la Figura H.5.c. En esta gráfica se representa simultáneamente la desviación estándar de la medida de actividad (línea azul) y la desviación estándar del modelo (línea roja). Se muestra el valor real de las desviaciones (líneas por encima de 0) y los valores negativos (líneas por debajo de 0). Estas dos variables junto con la actividad filtrada (línea negra en la Figura H.5.b) son las variables de entrada a la fase III del método propuesto (detección de comportamiento anómalo). Como ya se ha descrito, a través de las desviaciones calculadas se construye una función de pertenencia Fuzzy que proporciona para cada muestra de actividad física medida, un valor de anomalía (Figura H.6.a). Estos valores son filtrados en este ejemplo (Figura H.6.b) para tener a la salida todos los comportamientos anómalos detectados ($V = 0$) y de duración superior a 1 hora (60 muestras, $N_L = 60$).



a) Modelo de comportamiento y medida de actividad del día evaluable

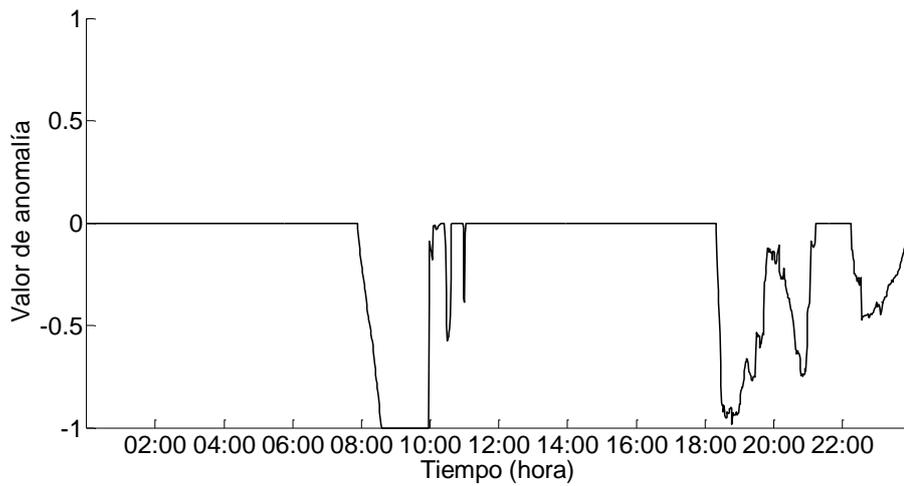


b) Diferencia de actividad calculada y filtrada

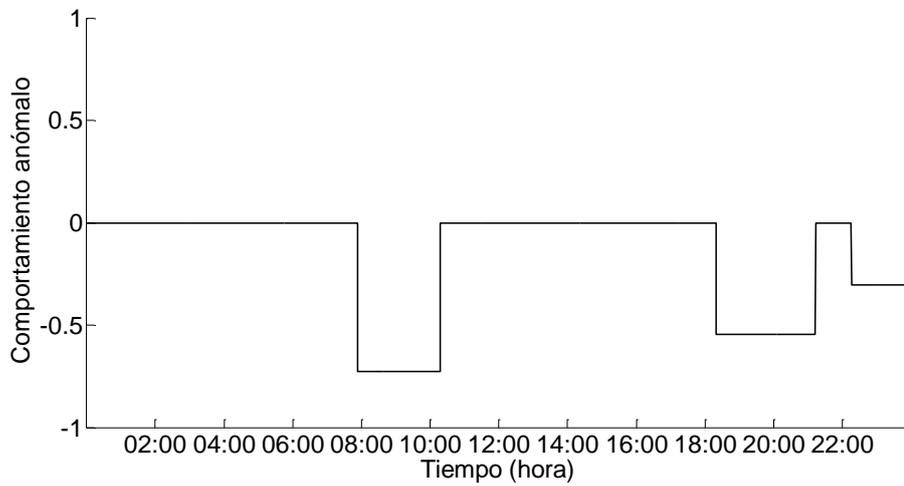


c) Diferencia de actividad y valores de desviación estándar para cada muestra de actividad

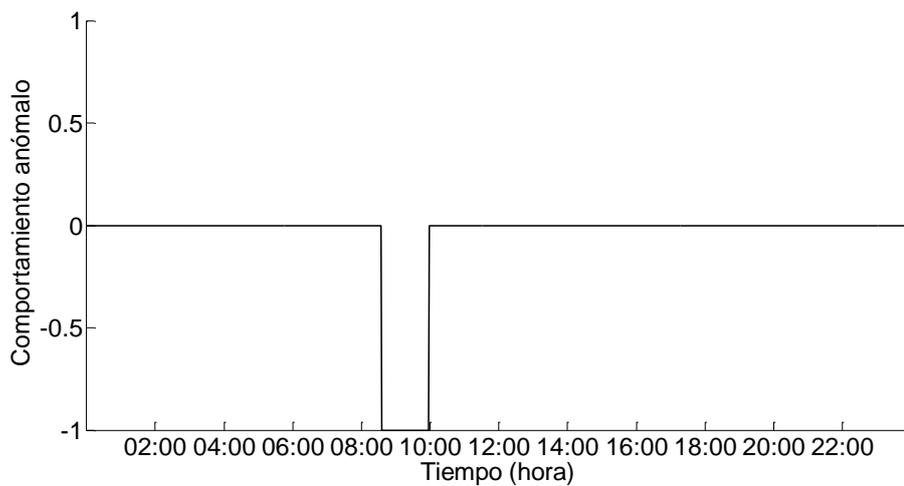
Figura H.5. Ejemplo de aplicación del método de detección de anomalías (I)



a) Valores de anomalía dados por la función Fuzzy de valoración



b) Comportamiento anómalo detectado con $N_L = 60$ y $V = 0$



c) Comportamiento anómalo detectado con $N_L = 60$ y $V = 1$

Figura H.6. Ejemplo de aplicación del método de detección de anomalías (II)

La salida del método de detección de anomalías, mostrada en la Figura H.6.b, es una variable continua de valores comprendidos entre -1 y 1 que describe el comportamiento anómalo detectado. En el caso del ejemplo de la Figura H.6, todos los comportamientos anómalos detectados (Figura H.6.b) son debidos a decrementos de la actividad del usuario.

Modificando los valores V y N_L del método de análisis, la salida proporcionada por el método varía considerablemente. La variación de estos valores permite la detección de comportamiento anómalo de forma adaptativa al contexto. Por ejemplo, manteniendo el tamaño del comportamiento anómalo detectado ($N_L = 60$) pero modificando el parámetro V del valor 0 (sin filtrado por valor) al valor 1 (se tendrá a la salida solamente el comportamiento anómalo valorado con el valor máximo), se obtiene la variable de comportamiento anómalo mostrada en la Figura H.6.c En esta nueva variable, solamente se identificaría el comportamiento anómalo ocasionado por el cambio de hora de levantarse.

APÉNDICE I.

DETALLE DE LOS RESULTADOS DE LA VALIDACIÓN

I.1. LISTA DE USUARIOS

Este apéndice recoge los detalles sobre los 4 usuarios participantes en la etapa de experimentación llevada a cabo en el Capítulo 7 para la validación global del método propuesto. Cada usuario fue monitorizado durante 90 días. Los resultados obtenidos durante este experimento son descritos y analizados en el Capítulo 7. La Tabla I.1 muestra para cada uno de los usuarios involucrados en la experimentación su sexo, edad, peso, estatura e índice de masa corporal (BMI). El perfil de actividad del usuario está relacionado con el tipo de ejercicio físico realizado durante el periodo de monitorización.

Tabla I.1. Lista de usuarios involucrados en la validación global del método

Nº	Sexo	Edad	Peso (Kg)	Estatura (m)	BMI	Duración experimento (días)	Perfil de actividad
1	V	71	75,0	1,60	29,29	90	Bajo
2	M	63	54,0	1,72	18,25	90	Bajo
3	V	64	77,0	1,76	24,86	90	Medio
4	V	78	70,0	1,65	25,71	90	Medio

I.2. COMPORTAMIENTOS ANÓMALOS DETECTADOS

Las Tabla I.2, Tabla I.3, Tabla I.4 y Tabla I.5 muestran los comportamientos anómalos detectados para cada uno de los usuarios monitorizados. Por cada anomalía de comportamiento detectada se identifica el número de día de monitorización, el día de la semana, el instante inicial y final y el valor de anomalía dado por el método propuesto. Junto con esta información se incluye el nivel de actividad reportado por el usuario para el día y tramo horario en el que se produce el comportamiento anómalo y el nivel de actividad reportado durante ese tramo horario para los $N_r = 5$ días anteriores utilizados en el método de identificación de rutinas. Estos últimos valores son obtenidos siguiendo el procedimiento descrito en el Capítulo 4 y calculando el valor medio para el tramo horario en el que se detecta el comportamiento anómalo.

Tabla I.2. Comportamientos anómalos detectados para el usuario 1

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
13	Miércoles	18:40	19:41	-1	1,0	1,8
	<i>Hace mal tiempo y el usuario no puede salir a pasear</i>					
15	Viernes	08:30	10:04	+1	2,0	0,8
	<i>El usuario se despierta antes que de costumbre para acudir al médico</i>					
17	Lunes	18:20	19:20	-1	0,0	1,8
	<i>El usuario quitó el reloj durante una hora para hacer limpieza general en casa</i>					
18	Martes	17:04	19:00	+1	2,0	1,6
	<i>Después de varias tardes en casa, hace buen tiempo y el usuario sale a pasear</i>					
21	Viernes	09:55	13:27	-1	0,0	2,0
	<i>El usuario no se coloca el dispositivo hasta la hora de la comida</i>					
22	Sábado	09:58	11:42	-1	0,0	1,6
	<i>El usuario se despierta más tarde que de costumbre</i>					
26	Miércoles	09:08	10:51	+1	2,0	0,0
	<i>El usuario madruga más de lo habitual</i>					
31	Martes	20:13	21:19	-1	1,0	2,0
	<i>El usuario se encuentra mal y se acuesta antes de lo que acostumbra</i>					
37	Miércoles	09:15	10:31	+1	2,0	1,4
	<i>El usuario madruga más que los últimos días</i>					
40	Sábado	12:04	13:38	-1	0,5	2,0
	<i>El usuario se encuentra mal y no sale a hacer la compra</i>					
43	Martes	11:11	12:17	-1	0,0	2,0
	<i>Al usuario se le olvida colocar el dispositivo hasta media mañana</i>					
46	Viernes	09:28	10:52	+1	2,0	1,0
	<i>El usuario sale antes a hacer la compra porque a media mañana tiene que acudir al médico</i>					
48	Domingo	13:47	15:15	-1	0,0	2,0
	<i>El usuario quita el dispositivo para fregar y se le olvida ponerlo hasta después de comer</i>					
50	Martes	09:25	11:04	-1	0,0	1,6
	<i>El usuario se levanta más tarde de lo habitual</i>					

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
54	Sábado	10:43	11:45	-1	0,0	1,6
	<i>El usuario se coloca el dispositivo de monitorización a media mañana</i>					
58	Viernes	13:50	15:09	-1	0,0	1,8
	<i>El usuario quita el dispositivo para fregar y no vuelve a ponerlo hasta que termina de comer</i>					
61	Lunes	18:31	21:13	-1	0,0	1,8
	<i>El usuario quita el dispositivo para limpiar y no vuelve a ponerlo hasta después de cenar</i>					
65	Domingo	09:00	11:21	-1	1,0	2,0
	<i>El usuario se levanta más tarde de lo habitual</i>					
67	Martes	17:35	19:33	-1	1,0	2,0
	<i>Hace mal tiempo y el usuario se queda en casa en lugar de salir a pasear</i>					
70	Viernes	08:25	09:37	+1	2,0	0,7
	<i>El usuario madruga para ir al médico</i>					
72	Lunes	07:34	09:03	+1	2,0	0,8
	<i>El usuario madruga para ir al médico</i>					
73	Martes	17:31	18:31	-1	1,0	1,8
	<i>Hace mal tiempo y el usuario no sale a pasear</i>					
85	Domingo	10:06	12:35	-1	0,7	2,0
	<i>El usuario se levanta más tarde de lo que acostumbró durante los últimos días</i>					
87	Martes	17:18	18:34	-1	1,0	1,6
	<i>Hace mal tiempo y el usuario no puede salir a pasear a media tarde</i>					
89	Jueves	18:36	19:39	+1	2,0	1,4
	<i>Después de varios días con mal tiempo, el usuario vuelve de nuevo a salir a pasear por las tardes</i>					

Tabla I.3. Comportamientos anómalos detectados para el usuario 2

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
09	Sábado	08:21	09:25	-1	0,0	2,0
	<i>El fin de semana, el usuario se levanta más tarde que por semana</i>					
09	Sábado	15:14	17:03	-1	0,5	1,9
	<i>El fin de semana, el usuario se levanta más tarde que por semana</i>					
09	Sábado	21:31	22:46	-1	0,5	1,6
	<i>Finaliza su actividad antes de lo habitual</i>					
15	Domingo	11:16	12:16	+1	3,0	2,0
	<i>El usuario hace una ruta de monte. Por semana, a esta hora está en la oficina</i>					
21	Sábado	08:53	16:15	-1	0,0	2,0
	<i>Al usuario se le olvida ponerse el reloj hasta por la tarde</i>					
22	Domingo	10:40	12:42	+1	3,0	2,0
	<i>El usuario hace una ruta de monte. Por semana, a esta hora está en la oficina</i>					
33	Domingo	20:31	21:34	-1	0,0	2,0
	<i>Finaliza su actividad antes de lo habitual</i>					
39	Lunes	08:43	10:08	-1	0,0	2,0

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
	<i>Se trata de un día festivo. Se levanta más tarde de lo habitual</i>					
50	Sábado	08:28	09:36	-1	0,5	2,0
	<i>El fin de semana, el usuario se levanta más tarde que por semana</i>					
51	Domingo	10:29	12:16	+1	3,0	2,0
	<i>El usuario hace una ruta de monte. A esta hora, por semana, suele estar en la oficina y no realizando actividades deportivas</i>					
55	Sábado	08:23	09:36	-1	0,0	1,9
	<i>El fin de semana, el usuario se levanta más tarde que por semana</i>					
59	Miércoles	15:40	16:42	-1	2,0	1,8
	<i>A la hora de la comida, el usuario acompaña a su madre al médico</i>					
60	Jueves	21:24	22:38	-1	0,0	1,8
	<i>El usuario se acuesta antes de lo que acostumbra</i>					
62	Sábado	08:33	09:36	-1	0,0	2,0
	<i>El fin de semana, el usuario se levanta más tarde que por semana</i>					
66	Viernes	12:23	13:49	+1	3,0	2,0
	<i>El usuario hace una ruta de monte. Se trata de un día festivo. Su rutina a esta hora es estar en la oficina</i>					
84	Sábado	16:35	17:46	+1	3,0	1,8
	<i>El usuario hace una ruta de monte. Por semana, a esta hora está en la oficina</i>					

Tabla I.4. Comportamientos anómalos detectados para el usuario 3

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
11	Martes	09:16	11:37	-1	1,3	2,5
	<i>El usuario se encuentra mal y no sale a pasear</i>					
15	Sábado	08:40	09:50	-1	2,0	2,6
	<i>El usuario sale a pasear pero con menor intensidad de lo que acostumbra</i>					
20	Viernes	10:03	11:11	-1	2,0	2,4
	<i>El usuario se encuentra mal y termina de pasear antes de lo que acostumbra</i>					
28	Sábado	08:08	09:37	-1	1,0	2,5
	<i>Hace mal tiempo y en lugar de salir a pasear, el usuario se queda en casa</i>					
45	Sábado	08:35	09:57	-1	2,0	3,0
	<i>El usuario pasea con menor intensidad ya que se encuentra mal</i>					
62	Sábado	08:44	10:01	-1	0,0	2,4
	<i>El usuario se levanta más tarde de lo habitual y no sale a pasear. Está enfermo</i>					
63	Domingo	09:20	11:26	-1	0,0	2,1
	<i>El usuario se levanta más tarde de lo habitual y no sale a pasear. Está enfermo</i>					
69	Sábado	08:44	10:39	-1	2,0	3,0
	<i>El usuario sale a pasear pero con menor intensidad de lo que acostumbra</i>					
73	Viernes	17:00	18:01	+1	2,0	1,0
	<i>Hace muy buen tiempo y el usuario decide salir a caminar junto con unos amigos</i>					
81	Sábado	10:07	11:14	-1	2,0	2,8
	<i>El usuario sale a pasear pero menos tiempo de lo que acostumbra</i>					

Tabla I.5. Comportamientos anómalos detectados para el usuario 4

Fecha	Día	Tiempo inicial	Tiempo final	Valor anomalía	Nivel actividad día anómalo	Nivel actividad días anteriores
16	Sábado	11:45	12:57	-1	0,0	3,0
	<i>Comienza su actividad más tarde de lo habitual. Se encuentra mal y decide no salir a pasear</i>					
29	Domingo	10:48	11:53	-1	0,0	3,0
	<i>Comienza su actividad más tarde de lo habitual. Se despierta y se le olvida poner el dispositivo hasta alrededor de las 12 de la mañana tras haber dado su paseo rutinario</i>					
34	Viernes	11:29	12:31	+1	3,0	1,6
	<i>Durante la última semana, el usuario no ha salido a pasear por la mañana porque ha estado enfermo. Hoy comienza de nuevo a salir a pasear</i>					
48	Martes	12:02	13:08	+1	3,0	1,4
	<i>El usuario lleva unos días sin salir a pasear por la mañana y decide retomar su rutina</i>					
50	Jueves	16:44	17:46	+1	3,0	2,0
	<i>El usuario sale a pasear por la tarde en lugar de por la mañana</i>					
60	Martes	16:59	18:05	-1	0,0	2,0
	<i>El usuario quita el dispositivo para hacer tareas de bricolaje en casa</i>					