ROBOTS IN ASSISTED LIVING ENVIRONMENTS

UNOBTRUSIVE, EFFICIENT, RELIABLE AND MODULAR SOLUTIONS FOR INDEPENDENT AGEING

DELIVERABLE 3.1
Conceptual architecture for sensing methods and sensor data sharing I

Dissemination Level	**Public**
Due Date of Deliverable	Project Month 6, September 2016
Actual Submission Date	30 September 2015
Work Package	WP3, *Modular conceptual home architecture design and ICT method development for efficient, robust and flexible elder monitoring and caring*
Task	T3.1, *Conceptual architecture for sensing methods and sensor data sharing*
Lead Beneficiary	S&C
Contributing beneficiaries	NCSR-D, TWG, RUB, AVN
Type	Report
Status	Submitted
Version	Final

Abstract

Architecture document, pertaining to the information dependencies and interoperability between RADIO components for satisfying the medical information requirements set in WP2. This document guides the development of recognition methods in Task 3.2.

History and Contributors

Ver	Date	Description	Contributors
01	10 Jul 2015	First draft, establishing document structure and work allocation.	S&C and NCSR-D
02	20 Jul 2015	Current Smart Home capabilities (Section 3)	S&C
03	18 Aug 2015	Analysis of interRAI items to identify relevant recognition methods (Section 2)	NCSR-D
04	3 Sep 2015	Brief introduction to the recognition methods and their maturity (Section 2)	NCSR-D, TWG, AVN
05	10 Sep 2015	Analysis of communication requirements and development of communication architecture (Section 4)	TWG, RUB
06	15 Sep 2015	Identification of Smart Home elements that can provide additional data for interRAI assessment (Section 3)	S&C
07	25 Sep 2015	More information about image analysis and localization methods (Section 2)	RUB
08	30 Sep 2015	Internal review comments and corrections	RUB
Fin	30 Sep 2015	Final document preparation and submission	NCSR-D

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Purpose and Scope

This deliverable is the first version of the *conceptual architecture* that specifies the information dependencies and interoperability between RADIO components for satisfying the medical information requirements set in WP2.

Within the scope of this document is:

- To identify the recognition methods that will be extended and adapted in the context of Task 3.2 in order to extract the interRAI assessment items that are required by D2.2.
- By studying the information dependencies between these components, to specify the conceptual architecture of the RADIO Home system that integrates the processing elements executing the recognition components above with the sensing elements (including smart home and robot-mounted sensors).

Outside the scope of this document is:

- To specify the physical architecture, such as whether components will execute on FPGA boards or general-purpose computers. This will be dealt in Task 4.1.
- To specify the architecture (either conceptual or physical) of the communication between the RADIO Home and other nodes of the RADIO ecosystem, such as cloud storage components and components meant to be used by hospital personnel or informal care-givers. This will be dealt in Task 5.1.

1.2 Approach

This deliverable documents work in Task 3.1, which analyses the user requirements collected by work in WP2 in order to develop the system architecture. The work plan in this task includes:

- Identifying recognition methods that can cover requirements on collecting information.
- Studying the data interdependencies between the different recognition methods in order to specify what intermediate results can be shared. This leads to providing technical requirements regarding data processing and data transfer.
- The analysis above can also identify alternatives (data and processing) for extracting the same information item. This can be used to balance between robustness and ethical requirements. For example, suppose bathroom ADLs can be recognized using only acoustic data or (more robustly) using a combination of acoustic and visual data. This deliverable will describe the alternatives in order to provide input to the relevant ethical discussion in WP2.

S&C leads this task and provides knowledge and experience with current smart home systems to match requirements from WP2, and the information exchange between the smart home system and the robot. **NCSR-D**, **TWG**, and **AVN** carry out the analysis that identifies appropriate recognition methods for the interRAI assessment items. **TWG** and **RUB** explore different approaches focusing on WSN and short range wireless communication technologies that can be exploited in RADIO.

The definition of the conceptual architecture in Task 3.1 has also allowed the consortium to refine the approach to WP3 as a whole and to establish the work plan, allocating work on specific recognition methods to the beneficiaries of WP3 (Table 1).

Table 1: WP3 work plan

Technology	Beneficiary	Sub-task
Audio event recognition; cf. Sect. 2.2.1	NCSR-D	Development of the RADIO audio acquisition and feature extraction platform, in C/C++, assuming pyAudioAnalysis as a starting point.
	NCSR-D	Experimenting with the SVM classifier in pyAudioAnalysis and evaluation on RADIO use cases.
	TWG	Experimenting with the Theodorou et al. (2014) method and with various open source toolkits and evaluation on RADIO use cases.
Speaker identification and verification; cf. Sect. 2.2.2	NCSR-D	Experimenting with pyAudioAnalysis and Spear and evaluation on RADIO use cases.
Acoustic mood recognition; cf. Sect. 2.2.3	NCSR-D	Experimenting with pyAudioAnalysis and evaluation on RADIO use cases.
	TWG	Experiments with short-time analysis of audio recordings, the Kostoulas et al. architecture, and open source toolkit. Evaluation on RADIO use cases.
Automatic speech recognition; cf. Sect. 2.2.4	NCSR-D	Integrating and training the CMU Sphinx ASR system.
Human motion analysis; cf. Sect. 2.2.5	NCSR-D	Experiments with depth and colour information analysis methods implemented in OpenCV and evaluation on RADIO use cases.
	NCSR-D	Implementation of the Varvadoukas et al. (2012) method in Python or C/C++ to allow integration. Extensions and evaluation on RADIO use cases.
Silhouette analysis; cf. Sect. 2.2.6	AVN	Experiments with methods for detecting level of activity, location, and posture and evaluation on RADIO use cases.
Face detection, recognition, expression analysis; cf. Sect. 2.2.7	NCSR-D	Experiments with the OpenCV implementation of the Viola & Jones algorithm and evaluation on RADIO use cases.
	NCSR-D	Experiments with fusing visual information with other modalities and evaluation on RADIO use cases.
	NCSR-D	Experiments with SHORE and evaluation on RADIO use cases.
Clothes recognition; cf. Sect. 2.2.8	NCSR-D	Developing the visual analytics for detecting change in clothing.
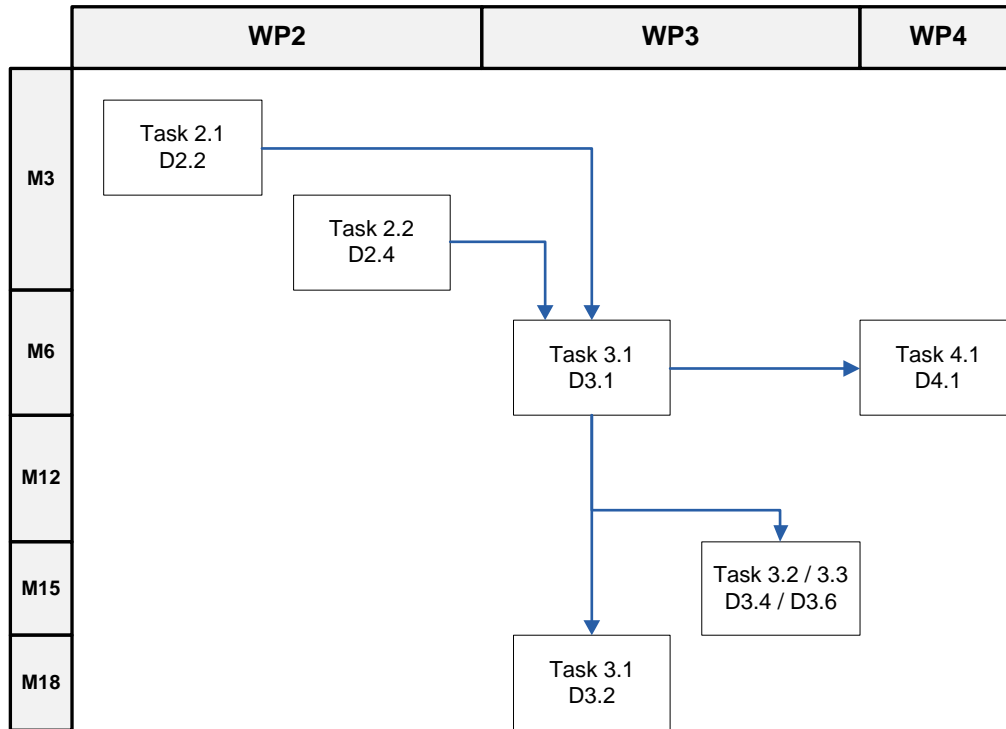
Figure 1: Relation to other Work Packages and Deliverables

1.3   Relation to other Work Packages and Deliverables

This document is the first in a series of closely related deliverables. This initial architecture and its next edition (D3.2, M18) are used to drive development in Task 3.2. From M19 until M24 the architecture is maintained as a *living document* updated to record adjustments necessitated as development in Task 3.2 progresses. The final version (D3.3, M24) documents the architecture and interfacing of the final data analysis methods (D3.5).

This deliverable is prepared based on the medical and ethical requirements set in D2.2 *Early detection methods and relevant system requirements* and D2.4 *Actual and perceived privacy considerations and ethical requirements*, respectively. The conceptual architecture developed in this deliverable is used by Task 4.1 in order to prepare the first version of the physical architecture, D4.1 *Architecture for extending smart homes with robotic platforms*.

2 DETERMINING INTERRAI ITEMS THROUGH SENSOR DATA

This chapter specifies the sensor data analysis technologies that will be developed in T3.2 and T3.4.

2.1 Recognition of interRAI Items

We assume as a starting point the interRAI assessment items that have been marked as relevant to RADIO in D2.2. Specifically, we assume the items marked as "Observation" in D2.2, since these are assessed by observing the user. Other interRAI items are assessed by interviewing the user or by accessing the medical record, and are outside the scope of work in WP3.

For each of the observation items, we provide an initial estimation of (a) the kind of data and processing over this data that can be used to assess the item; and (b) the maturity of the technologies involved in this processing.

These estimations are collected in Tables 2, 3 and 4, following the way interRAI organizes assessment items in ADL items, cognitive and mood items, and health conditions and continence items. For each item, we note the data modalities (visual, audio, 3D/range) that can be used for its recognition, the analysis technology that will be applied to this data, and the level of maturity of these technologies. The technologies are discussed in more detail in Section 2.2 below, but, in brief, they revolve around:

- Pattern recognition over visual (including 3D vision), range, and acoustic sensor data
- Extracting mood and other behaviour features from speech content.
- It is also foreseen that mood will be extracted from textual content and on-line social networking interactions. This is not treated in this version of the architecture, but will be included in the next version (D3.2), as it also depends on specifications work to be carried out in WP5.

An estimation of the level of maturity of the specific technologies is given in the Technology Readiness column of Tables 2, 3 and 4 as follows:

1. It is unclear how this recognition can be achieved, and novel algorithmic design is needed. Although some initial experiments might be carried out, usable and robust outcomes are beyond the scope of the RADIO project.
2. There is a clear path to achieving this by adapting existing algorithms and methods. The compilation of annotated test data and extensive experimentation will be needed, but usable and robust outcomes are expected for most of these items by project's end.
3. The methods for achieving the recognition of these items are already developed. Minor testing and tuning is needed. Usable and robust outcomes for all of these items will be integrated in the first RADIO prototype.

Table 2 ADL-related data extraction methods

Assessment Item	Name	Technology	Technology Readiness	Comments
ADL Self Performance	Bathing, personal hygiene	Audio event recognition	3	Annotated audio recordings are needed.
		Audio-visual event recognition	2	If visual sensor in bathroom is acceptable
	Dressing upper body; dressing lower body	Clothes recognition	2	Visual sensor
	Walking, locomotion	Human pattern recognition (HPR) in range data and face identification	2	Assuming the scene is not crowded, face identification is used to identify user and HPR to characterize walking patterns.
	Toilet transfer	Visual event recognition Simple proximity sensor	2	If visual sensor in bathroom is acceptable. Proximity sensor can give simple on/off detection.
	Toilet use	Audio event recognition	1	Annotated audio recordings are needed.
	Bed mobility	Visual event recognition	2	
	Eating	Visual event recognition	2	This could be difficult, especially if it is necessary to distinguish from simply sitting at the table and fumbling with food. Annotated samples are needed.
Locomotion and Walking	Timed 4-meter walk; distance walked	Human pattern recognition (HPR) in range data and face identification	2	Face identification is needed to identify user. HPR provides needed measurements.
	Distance wheeled self	*N/A, as ability to walk is required (cf. inclusion criteria, Sect 2, D2.2)*		
Activity Level	Total hours of exercise or physical activity in last 3 days	Visual	2	Assuming it is not necessary to classify to different exercise activity types.
	Number of days went out of the house in last 3 days	Visual	2	Detecting exit and entrance. Simple post processing can separate substantial exits from going just out of the door and back again.

Table continues from previous page.

Assessment Item	Name	Technology	Technology Readiness	Comments
Activity Preferences and Involvement	Cards, games, or puzzles; crafts or arts; reading, writing, or crossword puzzles	Audio-visual event recognition	1	Annotated examples in a controlled environment are needed.
	Computer activity	Audio-visual event recognition	1	Annotated recordings are needed. Will also consider monitoring the computer.
	Conversing or talking on the phone	Audio-visual event recognition	3	Annotated recordings are needed. Visual not needed if it is not required to distinguish between conversing and phone.
	Dancing	Audio-visual event recognition	2	Annotated recordings are needed.
	Exercise or sports	Audio-visual event recognition	2	Annotated recordings are needed.
	Gardening or plants	*N/A, RADIO operates indoors*		
	Helping others	Audio-visual event recognition	1	Annotated recordings are needed.
	Music or singing	Audio-visual event recognition	3	Annotated recordings are needed.
	Pets	*N/A, presence of pets is not considered in the RADIO experiment environment*		
	Spiritual or religious activities	Visual event recognition, automatic speech recognition	1	Examples and characteristic poses, keywords, key-phrases are needed.
	Trips/shopping	*N/A, RADIO operates indoors*		
	Walking or wheeling outdoors	*N/A, RADIO operates indoors*		
	Watching TV or listening to radio	Audio-visual event recognition	3	Detecting if TV or radio is on, detecting if end-user is actually watching TV.
	Discussing and reminiscing about life	Automatic speech recognition	1	Examples and characteristic keywords and key-phrases are needed.
Time Asleep During Day	Time Asleep During Day	Audio-visual event recognition	2	Sensors in bedroom. It will be harder to consider sleeping in couch or on chair (maturity 1).

Table 3. Cognitive and Mood Items data extraction methods

Assessment Item	Name	Technology	Technology Readiness	Comments
Cognitive skills for daily decision making				Need more details: this item is high-level and needs more detailed description
Memory/ recall ability	Procedural memory			Too general. Will be re-assessed in D3.2.
	Situational memory			
Periodic Disordered Thinking Or Awareness	Episodes of disorganized speech	Audio	1	Annotated recordings are needed.
Acute Change in Mental Status from Person's Usual Functioning				Too general. Need more details
Indicators of possible depressed anxious or sad mood	Made negative statements	Audio (Speech)	2	Detailed examples of speech (languages?)
	Persistent anger with self or others	Audio	2	Audio recordings and annotation of affect status are needed.
	Expressions, including non-verbal, of what appear to be unrealistic fears	Audio-visual event recognition	1	Examples AND training data (annotated samples)
	Repetitive health complaints; repetitive anxious complaints and (non-health related) concerns	Automatic speech recognition	2	Detailed examples and characteristic keywords are needed.
	Sad, pained, or worried facial expressions	Facial expression analysis	1	Annotated samples. It is extremely difficult to get reliable estimations, unless calibration on each individual person can be applied.
	Crying, tearfulness	Audio event recognition	2	Annotated samples are needed.
	Recurrent statements that something terrible is about to happen	Automatic speech recognition	2	Detailed examples and characteristic keywords are needed.
	Withdrawal from activities of interest; Reduced social interactions		–	Too general. Will be re-assessed in D3.2.
	Expressions, including non-verbal, of a lack of pleasure in life	Automatic speech recognition	2	Detailed examples and characteristic keywords are needed. Non-verbal expressions cannot be reliably recognized (level: 1)

Behaviour Symptoms	Wandering	Visual	2	Need annotated samples and advanced post-processing to fuse with other detected events
	Verbal abuse	Automatic speech recognition	2	Characteristic keywords are needed.
	Physical abuse	Audio-visual event recognition	1	Need annotated samples
	Socially inappropriate or disruptive behaviour	Automatic speech recognition	2	Detailed examples and characteristic keywords are needed.
	Inappropriate public sexual behaviour or public disrobing; Resists care	Audio-visual event recognition	1	Only if "public" is understood as "at home but in the presence of others." Will be re-assessed in D3.2.

Table 4. Health Conditions and Continence items data extraction methods

Assessment Item	Name	Technology	Technology Readiness	Comments
Falls	Falls	Visual event recognition	2	-
Health Condition Frequency	Difficult or unable to move self to standing position unassisted; to turn self around and face the opposite direction when standing	Visual event recognition	1	Annotated samples. One foreseen difficulty is to understand when there is a will to stand-up. Could count e.g. when is standing up but cannot know how many times tried and failed.
	Dizziness			It is unclear how this can be observed.
	Unsteady gait	Visual	2	Annotated recordings are needed.
	Difficulty clearing airway secretion; Constipation; Diarrhoea	Audio event recognition	1	Annotated audio recordings are needed.
	Vomiting	Audio-visual event recognition	1	Annotated recordings are needed.
	Difficulty falling asleep or staying asleep; waking up too early; restlessness; non-restful sleep; Too much sleep	Audio-visual event recognition	2	Visual sensor in bedroom
	Aspiration, coughing while eating	Audio event recognition	2	Depends on ability to recognize eating ADL (cf. above)
Dyspnea	Dyspnea	Audio and Visual for activities	2	Need annotated samples & examples
Fatigue	Fatigue		1	Unclear how this can be observed.
Pain symptoms	Frequency with which person complains or shows evidence of pain (including grimacing, teeth clenching, moaning, withdrawal when touched, or other nonverbal signs suggesting pain)	Audio-visual event recognition	1	Annotated recordings are needed.
	Consistency of pain	Audio-visual event recognition	1	Annotated recordings are needed.
Continence	Bladder continence			Will be revisited in D3.2

2.2 Methods for Extracting Required Clinical Information

2.2.1 Audio event recognition

Audio event recognition will be based on multiple classifiers that receive features from a single audio acquisition and feature extraction platform. The platform records and segments audio and then processes it to extract all the features needed by all classifiers for the various events. This platform will be based on the pyAudioAnalysis Toolkit.[1] The more complex setup that fuses multiple microphones and multiple instances of the platform will be dealt with in subsequent versions of the RADIO architecture. Alternative pre-processing and feature extraction methodologies will also be investigated, and their inclusion will also be dealt with in subsequent versions of the RADIO architecture.

For the classification itself, the consortium will experiment with the application of different approaches to indoors scenarios, including:

- The Support Vector Machine implementation in the pyAudioAnalysis Toolkit
- The Theodorou et al. (2014) method for robust recognition of indoors sounds of interest
- Methods implemented in various open source toolkits, such as Auditory Toobox[2], VoiceBox[3], OpenSmile[4], WEKA[5], and HTK[6].

The functionalities provided by these tools will appropriately be combined with other software tools/components developed within RADIO and adapted to the specifications and needs of the RADIO project. The latest trends in sound event recognition will be investigated for the needs of the RADIO project, such as hierarchical classification (Pellegrini et al., 2009; Niessen et al., 2013) and fusion architectures (Ntalampiras et al., 2008). In the case of hierarchical classification, the audio events are in general grouped depending on their sound similarity and classified in incrementally splitting pairs of sounds, while in the case of fusion architectures, late and early integration schemes can be followed for fusion on feature level or decision level respectively.

The development of the sound models will use a combination of existing publicly available audio databases, such as Soundbible[7], and audio data collected within the RADIO project.

2.2.2 Speaker identification and verification

The goal of this module is to identify or verify speaker identity for acoustic features. Towards this end, the Spear[8] open-source Python library will be adopted to train and use speaker verification and identification functionalities. In addition, other supervised and semi-supervised functionalities from pyAudioAnalysis will be used to achieve speaker recognition and/or speaker clustering. Audio features will be extracted from the feature extraction module described above.

2.2.3 Acoustic mood recognition

This method extracts supervised estimates for the user's mood based on acoustic (non-verbal) features. The overall processing pipeline is similar to that of even detection, comprising audio acquisition, feature extraction, and classification.

---

[1] https://github.com/tyiannak/pyAudioAnalysis
[2] https://engineering.purdue.edu/~malcolm/interval/1998-010
[3] http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html
[4] http://www.audeering.com/research/opensmile
[5] http://www.cs.waikato.ac.nz/ml/weka
[6] http://htk.eng.cam.ac.uk
[7] http://soundbible.com
[8] https://pypi.python.org/pypi/bob.spear/1.1.2

The consortium will experiment with and evaluate existing open-source speech and audio tools for audio pre-processing, feature extraction and classification applied to the detection of the mood of the user (affective modelling).

To detect negative affective status and mood, RADIO assumes as a starting point audio pre-processing and feature extraction tools that have been used for audio-based affective modelling, such as OpenSmile[4]. Typical architectures (Kostoulas et al., 2007; Kostoulas et al., 2012) with short-time analysis of audio recordings, feature extraction and classification will also be followed here, evaluating different classification algorithms[5].

2.2.4   Automatic speech recognition

Speech recognition will be based on existing open source tools. In particular, CMU Sphinx[9] will be trained in the RADIO-specific context conditions to prepare a keyword spotting recognizer.

2.2.5   Human motion analysis

RADIO will experiment with multiple methods for analysing visual and range data to detect humans in the environment and to extract motion-related observations, such as walking speed and walking patterns.

Visual methods to detect humans independently from external influences (such as lighting conditions or changing background) depend on the *histogram of oriented gradients (HOG)* features (Dalal and Triggs, 2005). It is often employed in computer vision scenarios for human detection since it provides a high success rate in real-time. The basic idea of HOG is that an accurate feature set for a human body is based on its silhouette. The flow of the algorithm to extract this feature set can be divided into six parts: luminance calculation, gradient computation, magnitude and theta computation, cell histogram formation and block normalization. If a coloured image serves as input to the algorithm, the luminosity of each pixel has to be calculated. Luminosity calculation is of course excluded when a grayscale image serves as input, as the grayscale values represent the luminosity. The resulting luminosity image is then used for vertical and horizontal gradient computation. The output of the gradient computation is a horizontal and a vertical gradient image. Out of these two images the gradient magnitude and the direction for each pixel can be calculated. With these two parameters, the cell histogram can be computed. The last step consists of normalization of these histograms which improves the invariance to illumination and shadowing. Locally normalized HOG descriptors significantly outperform existing feature descriptors. Finally, a classifier such as a support vector machine (SVM) or an AdaBoost classifier can be used to detect humans. The classifier decides based on the feature set if it is a human or not.

Depth (3D vision) and colour information analytics will be implemented based on the OpenCV[10] library of computer vision algorithms, to extract visual features and to get basic image manipulation functionalities. Supervised models will be trained to classify motions to predefined categories. This will be based on the mlpy[11] and scikit-learn[12] toolkits.

Range data from a laser range finder will also be used to recognize motion patterns. RADIO will assume as a starting point the Varvadoukas el at. (2012) method for recognizing human motion patterns in laser range data. In the context of RADIO, the current MATLAB prototype will be re-implemented in Python or C/C++ in order to allow integration. The method itself will also be extended in order to extract information about the walking quality.

The information from these modalities will be *fused* to increase robustness.

---

[9] http://cmusphinx.sourceforge.net

[10] http://opencv.org

[11] http://mlpy.sourceforge.net

[12] http://scikit-learn.org

Figure 2: Example of the silhouette extracted from an image

2.2.6 Silhouette analysis

A valuable approach used by many methods in human motion and posture analysis is silhouette analysis. Silhouette analysis restricts the detection algorithm to not use the entire full-detail visual representation as captured by a camera but to a binary classification of all pixels in the analysed image in two categories: background and foreground.

From the algorithmic point of view a number of steps is necessary in order to get to a valuable silhouette image (Figure 2):

- Step 1: Image binarization (sometimes region-based)
- Step 2: Removal of background
- Step 3: Removal of "false foreground" (erosion)
- Step 4: Removal of "false background" (dilation)
- Step 5: Labelling (in case of many persons in same room, needs post-processing)

In terms of detection capabilities, this approach can be used (alone or in combination with other methods) to detect a person's level of activity, location, and posture.

2.2.7 Face detection, recognition, expression analysis

Detection of a face is the first step in accurately detecting a person and, in co-operation with other methods, to provide most of the information needed to understand a person's location and direction of sight. Face detection is also a prerequisite for face recognition and facial expression analysis.

Despite extensive research in the last few years, the majority of available face detection systems are based on derivatives of the original work of Viola & Jones (2004). This face detection method belongs to the class of feature detection algorithms. With the help of cascaded rectangular Haar-like features, robust face detection can be performed. Since larger images require more processing time for facial feature detection, the input image for the face detection should be reduced as far as possible beforehand. Therefore, a detected face in the last image is stored and the region of interest should be analyzed first before resorting to the analysis of the full image. If a face is already detected in the image, the position and size of the rectangle surrounding the face is stored for future face detections. A simple face tracking method is to increase the size of the region of interest around the detected face. Then a slightly larger region of interest than optimal is analyzed, but with a great increase in the detection rate between several images with moving people. With the help of depth images, the image size can be permanently reduced through foreground isolation resulting in accelerated face detection. Foreground isolation is achieved through binarization and contour detection on the depth image. For contour detection, the algorithm by Suzuki and Abe (1985) is used. Whenever no initial information about the position of the face in the image exists, the face detection system only searches for faces in the foreground region of the infrared image. Due to this method, the whole image size is never actually processed.

The OpenCV implementation of the Viola & Jones algorithm will be assumed as a starting point. Information fusion methods can improve robustness by using information from the depth and range modalities.

To recognize a face however is a completely different problem. People's faces may vary significantly but there are also many similarities that could trick even the most robust algorithms. However, in the RADIO case we are allowed to use some assumptions that can significantly enhance reliability:

- The number of persons to recognize is quite small
- It is possible to train the system with all the faces it may encounter. If a proper automatic re-training step is added each time a new persons joins the group or some persons leaves, the reliability can be drastically improved.

Regarding expression analysis, the SHORE[13] recognition engine will be employed.

2.2.8 Clothes recognition

Visual information analytics will be adopted to detect change of clothes (as required by the respective ADLs). Towards this end, human detection and basic texture modelling will be used to classify detected clothes and detect possible changes.

2.3 Mobile Platform Navigation

One fundamental problem in robotics is *Simultaneous Localization and Mapping (SLAM)*, also known as Concurrent Mapping and Localization (CML) (Thrun et al., 2005). It arises when neither a map of the surrounding, nor the actual position of the robot is known. Several algorithms such as the SLAM with Extended Kalman Filters (EKF) and the SLAM with Particle Filters address this issue. The robot simultaneously creates a map of the environment and localizes itself relatively to this map. Within the RADIO project, a SLAM algorithm will be implemented in order to allow self-localization estimation of the robot within the smart home. Without this feature, the flexibility of the robot would be limited to a predefined map created before bringing the robot into operation. Based on the results of the SLAM algorithm, the robot is aware of its own position and can further use this information to navigate autonomously through the environment, supporting the unobtrusive monitoring of the patient. This goal creates new challenges for man-machine interaction, as the robot should never become an obstacle for the human.

Table 5: List of computational and bandwidth requirements of the data extraction methods.

Method	Computational Requirements	Bandwidth Requirements
Audio event recognition	Raspberry	16 KHz + DOA information (minor)
Speaker identification–verification	Raspberry	16 KHz + DOA information (minor)
Audio-based mood recognition	Raspberry	16 KHz + DOA information (minor)
Automatic Speech Recognition	PC	16 KHz + DOA information (minor)
Human motion analysis	PC	Depth: 5 fps, 640x480; RGB: 5 fps, 1280x1024
Face analysis	PC	Depth: 5 fps, 640x480; RGB: 5 fps, 1280x1024

Notes:
1. In the case of streaming, all audio-related methods share the same input stream.
2. Feature extraction is the most computationally demanding step, but can be shared by all audio methods.
3. All signal analysis methods listed above require no buffering of "historical" data, affording robustness to dropped frames since recognition is only impaired for the duration of the data transmission failure.

---

[13] http://www.iis.fraunhofer.de/en/ff/bsy/tech/bildanalyse/shore-gesichtsdetektion.html

Localization is very important for the cognition task since it provides the most relevant data regarding current distances to obstacles for safe navigation through the environment. The localization task depends on the data from the perception task. For localization within the SLAM algorithm, a particle filter can be chosen since it easily incorporates information from different sensor types, thus enabling sensor data fusion. The particle filter belongs to the family of Bayesian filters and models probability distributions discretely through particles which represent different system states. The initial state is an evenly distributed set of particles which then gets updated with the help of sensor information. The sensor information can be translated into a distribution function which is further used by the particle filter to calculate the weights of each particle resembling a system state. The weight calculation can be performed in several ways. An easy approach is to use the difference between the position determined by the sensor data and the particle position which serves as input to a normal distribution function.

The better the particle state resembles the distribution function, the higher the particle weight will be. The new particle set resembling the current robot state is determined through the particle weights by resampling particles with low weights. One of the main benefits of the particle filter is that it is considered to be an elastic algorithm (Schwiegelshohn and Hübner, 2014) meaning that a higher accuracy and therefore a better representation of the probability distribution can be achieved by greater number of particles, thus directly influencing the processing requirements. Additionally, a motion update model can be applied to the particle filter. This further increases localization accuracy as the robots current movement is taken into account by the particles.

2.4 Functional Requirements on Wireless Communication

Based on the recognition methods described immediately above, we set the functional requirements in Table 5 for the communication infrastructure.

References

Dalal, N. and Triggs, B. (2005). "Histograms of oriented gradients for human detection". Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005).

Kostoulas, T., Ganchev, T., Mporas, I., Fakotakis, N. (2007). Detection of Negative Emotional States in Real-World Scenario. In *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, Patras, Greece, 29–31 October, 2007, pp. 502-509.

Kostoulas, T., Mporas, I., Kocsis, O., Ganchev, T. Katsaounos, N., Santamaria, J.J., Jimenez-Murcia, S., Fernandez-Aranda, F., Fakotakis, N. (2012). Affective speech interface in serious games for supporting therapy of mental disorders. *Expert Systems with Applications* 39(12), 15 September 2012, pp. 11072–11079.

Niessen, M.E., Van Kasteren, T. L. M., Merentitis, A. (2013) Hierarchical Sound Event Detection. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA).

Ntalampiras, S., Potamitis, I., Fakotakis, N. (2008). Automatic recognition of urban soundscenes. New Directions in Intelligent Interactive Multimedia, pp. 147–153.

Pellegrini, T., Portêlo, J, Trancoso, I Abad, A, and Bugalho, M. (2009). Hierarchical clustering experiments for application to audio event detection. In *Proceedings of the 13th International Conference on Speech and Computer.*

Schwiegelshohn, F. and Hübner, M. (2014). "An Application Scenario for Dynamically Reconfigurable FPGAs". In Proc. ReCoSoC 2014, Montpeiller, France, May 2014.

Suzuki, S. and Abe, K. (1985). "Topological Structural Analysis of Digitized Binary Images by Border Following." CVGIP 30 1, pp 32-46.

Theodorou, T., Mporas, I., Fakotakis, N. (2014). Audio Feature Selection for Recognition of Non-Linguistic Vocalization Sounds. *Artificial Intelligence: Methods and Applications*, Lecture Notes in Computer Science, vol. 8445, pp. 395–405.

Thrun, S., Burgard, W., Fox, D. (2005). *Probabilistic Robotics.* The MIT Press.

Varvadoukas, T., Giotis, I., Konstantopoulos, S. (2012). Detecting Human Patterns in Laser Range Data. In *Proceedings of the 20th European Conference on AI (ECAI 2012).* Montpellier, France, 27–31 August 2012.

Viola, P. and Jones, M. J. (2004). Robust Real-Time Face Detection. *International Journal of Computer Vision* 57, pp. 137–154.

3 SMART HOME ARCHITECTURE

3.1 Introduction

The smart home architecture in RADIO has several requirements that need to be fulfilled. The product that RADIO will employ for creating the smart home solution is called *enControl™*. *enControl™* provides an initial solution for home automation, and will be enhanced and upgraded complementing other technologies and developments within RADIO that will form the final RADIO solution.

enControl™ is a white label B2B product by Sensing & Control Systems providing a smart home solution.

At a glance, *enControl™* has four main functions:

1) Comfort
2) Security
3) Energy Management
4) Automation

Users of *enControl™* are able to monitor and control remotely their homes & business using smartphones or any device running Web browsers, like tablets or PCs.

enControl™ can be divided into three main components, (i) home devices, (ii) IoT Platform and (iii) graphical user interfaces (as shown in Figure 10).

At home level, *enControl™* is composed by (i) end devices (sensors and actuators) and (ii) home controller (Gateway). The gateway supports communication using different protocols, but mostly uses ZWAVE radio technology for the communication of off-the-shelf end devices and home controller.

There are more than 300+ companies providing ZWAVE more than 1300 products[14] which enables *enControl™* to deliver the four main functionalities highlighted above, which are extended in following short list[15] (from the point of view of the information being triggered by the end devices):

1) Comfort
   a. Climate monitoring
   b. Climate control
   c. Temperature, Humidity, CO2 (etc…) levels
2) Security
   a. Detection of door/window opening
   b. Detection of movement
   c. Detection of Smoke
   d. Detection of CO
   e. Detection of water basement
3) Energy Management
   a. Energy consumption
   b. Energy control (switch on/off electricity, water, gas, etc…)
4) Automation
   a. Switch on/off appliances
   b. Switch on/off lights

---

[14] According to http://z-wavealliance.org
[15] The list does not pretend to cover all possibilities, the Reader should understand that the smart home solution can integrate any ZWAVE standard product, thus enabling the functionality delivered by a particular product. For full list of product, please visit ZWAVE alliance web page.

c. Open/Close doors, curtains, shutters

By exploiting or aggregating smart home information per se, and/or with other RADIO components information, insights on the activity being generated by sensors can provide means of inferring the end-users' activity The combination of smart home with other RADIO components provides an assisted living environment that does not rely on any intrusive devices, such as wearables attached to the end-user.

The home controller is a product from Sensing & Control, and complements the smart home functions delivered by the IoT platform.

The IoT platform contains the core of the smart home solution. It provides an open REST API enabling the home controllers to exchange information bidirectionally, based on synchronous or asynchronous actions triggered by IoT or end users through *enControl™* interfaces.

The IoT platform serves as well as main repository of information, keeps historical information about end devices data (information and status) as well as basic actions triggered by different actors, so users of the smart home solution know in real time who-when-what of actions monitored (for instance "the TV set has been switched on by Maria on 11 July 2015, 20:30 CET").

The open REST API is used by user interfaces through web clients and smartphone apps in order to present to end users the information being acquired from home end devices, and the action than can be triggered to them, so it encapsulates the smart home functions offered.

It is important to notice that *enControl™* clients (smart phone apps and Web interfaces, 3rd party services, etc.) interact with the smart home through the API, making internet connections required. It should be noted that experience with deployment of the solution is that this introduces no perceptible delay in executing actions by ZWAVE devices.

The main home controller target is to timely exchange information between devices and the IoT Platform (and thus with end users), and information being generated by the home controller itself. However, additionally provides three important functions:

1) Management of ZWAVE network and devices
2) Basic pre-processing capabilities (mainly energy calculations)
3) Local information repository in order to deal with temporal internet cut offs (so historical information is kept to be sent when internet connectivity is resumed).

The current smart home solution provides an intuitive and friendly user interface, which is similar for tablets and PCs (web interface) and smart phones.

Figures 3 through 9 show a representative sample of the smart home interfaces.
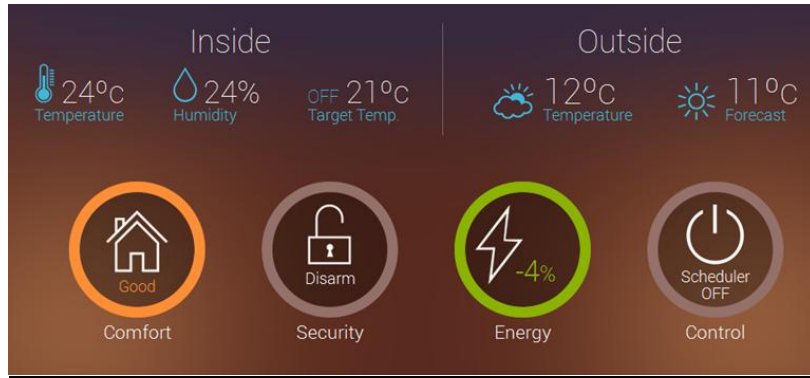
Figure 3: Control Panel
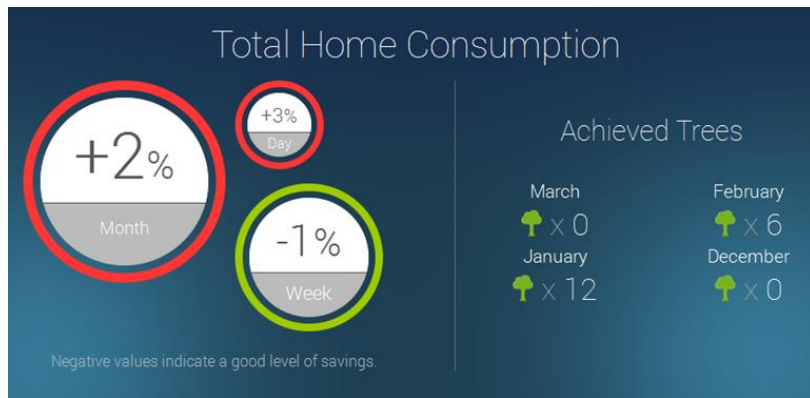

Figure 4: Climate control panel
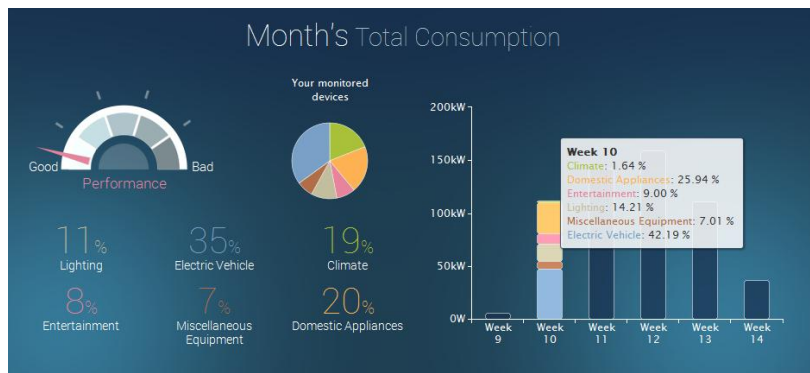

Figure 5:Energy consumption

Figure 6: Detailed energy consumption panel
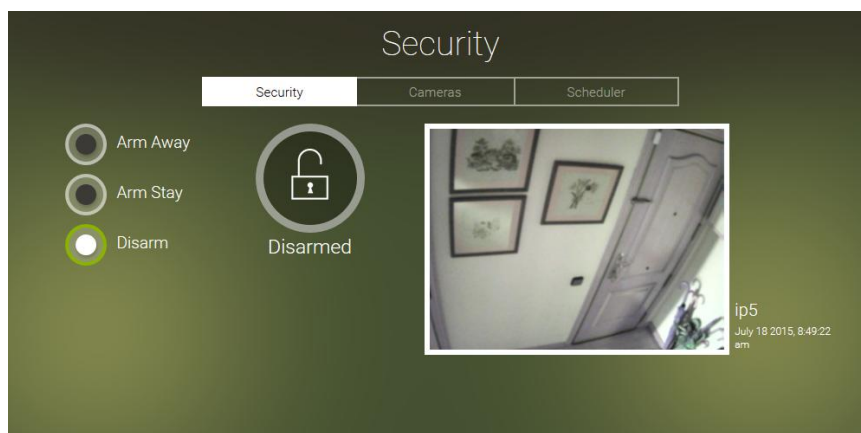
Figure 7: Control panel

Figure 8: Security panel

Figure 9: Smartphone interfaces

3.2 Home Automation Conceptual Architecture

Figure 10 represents *enControl's™* service architecture. Following up on the high level details about overall service provided above, this section focuses on those aspects of *enControl™* that facilitate the integration of the current *enControl™* version into the RADIO's AAL solution.

The IoT platform is able to connect, transport, process, analyse and manage data from the sensors to the real world to high-level applications and vice versa. It is able to operate in a booth wireless/wired network environments and supports different communication protocols. As a function summary:

- **Collect data**

  Collect data refers to the ability to retrieve and store information for further exploitation about all the parameters that are relevant to the system: messages, device status, commands, errors, exceptions.

- **Analyse data**

  The IoT Platform has processes in charge of analysing data and taking decisions based on the quality/importance/integrity of this data.

- **Data Aggregation/Data Fusion**

  Data aggregation refers to the ability to concatenate info from devices, for example when there is a mix of information between data from mash-up sensors.

  Data fusion, is the process of integration of multiple data and knowledge representing the same real-world object into a consistent, accurate, and useful representation

- **Translate data**

  This feature is related to the ability to transform raw data in bytes from the devices to high-level information.

- **System monitor**

  The IoT Platform is able to monitor the communication between devices and high level apps; and the overall status of the platform and device network. For instance, it is used to monitor key performance indicators like the number of messages sent per minute, number of exceptions/errors, number of devices connected to the platform, etc.

- **Transfer**

  The platform provides mechanisms to publish subscribe information to the queues exposed by the SDK

- **Audit information**

  In order to enable traceability of the functionality of all sub-systems, active logging is implemented within the software components. Log files are stored and available for analysis.

- **Secure components and communications**

  Communication between components and devices are secured

The current solution of *enControl™* provides means of interaction with home devices by accessing dedicated functions of the open API of the IoT platform. This implies that other ICT solutions (either at home like other RADIO components or remotely) willing to interact with *enControl™* must have internet connection.
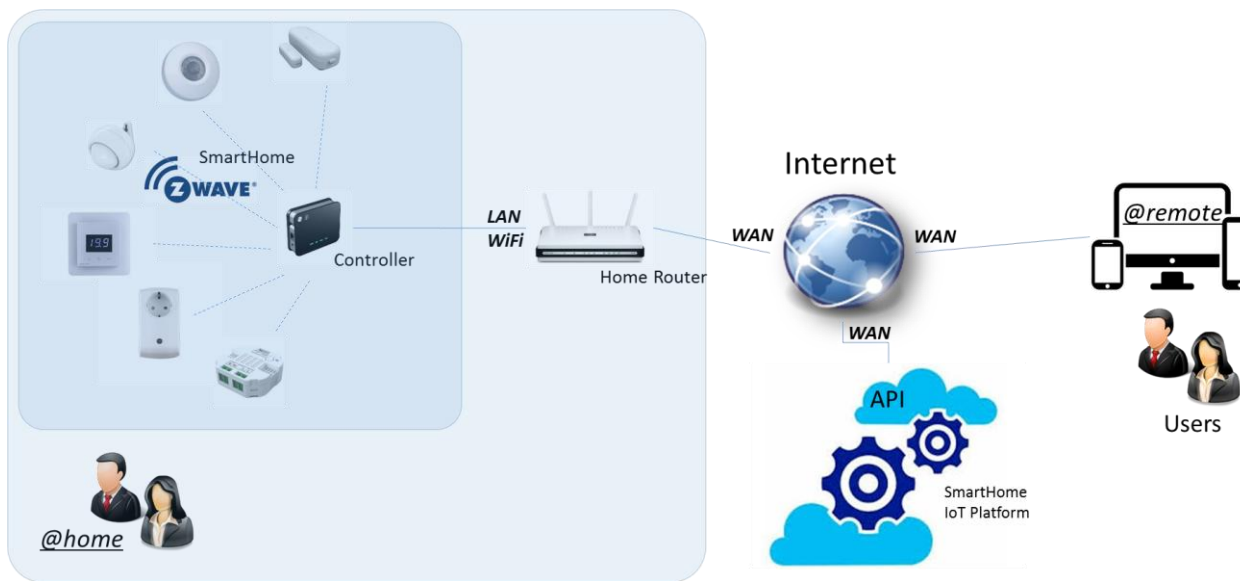
Figure 10: enControl™ *Service Architecture*

The API is divided into following main groups:

1) Authorisation and Authentication
   a. Functions used to grant access to different IoT resources and API functions
2) Devices
   a. Functions related to push and retrieve device status and data (including historical data)
3) Users
   a. Functions related to the management of users
4) Installations
   a. Functions related to the management of smart home installations
5) Monitor
   a. Functions related to the status monitoring of the IoT

On the other hand, the smart home controller has computational resources as well as physical interfaces that can be used to connect third-party transceivers so to enable connectivity via WiFi or BT for example. The controller is a compendium of Java programs, from HAL layer (devoted to ZWAVE communication, and other available HW transceivers) to the smart home application (controller part) running on top of a Linux-enabled controller. For RADIO, *enControl™* controller will be delivered running on Raspberry Pi.

Any additional code targeted to run on the home controller has to be decided to be included considering that it should affect the main functions currently running minimally, otherwise the responsiveness of the controller in front of potential user actions may be compromised leading to poor performance, which in turn may trigger potential rejection of the overall solution.

Table 6 summarizes the functionality from the smart home to be exploited by the overall RADIO solution (locally at home), and at which level this can be achieved and the expected effort (using a three level scale when not already available)

3.3   Smart Home contribution to recognition of interRAI Items

Section 2 provides a summary of InterRAI assessment items for different ADL items, cognitive and mood, conditions and continence items, which have been listed in D2.2. Section 2 provides a wide vision on different technologies (deployable at home) capable to provide means of assessment of the InterRAI items, in combinations with health records and interviews.

This section provides a link between the different capabilities of the smart home (based on sensors data and current functionalities) that can be used to complement the methods described in section 2.2.

The Smart Home solution continuously collects and stores data and status of different sensors and actions requested to actuators. The type of data collected directly depends on the off-the-shelf devices available on the market and deployed at home.

Table 7 provides an overview of the potential of smart home sensor data to complement the sensors and methods presented in Section 2.2. Further exploration will be done together in coordination with WP2 and other tasks in WP3 to take full profit of smart home service on this goal.

Table 6: Local Smart Home API functions vs. development effort

Smart Home API Function	Impact on ICT resource		
	IoT	Gateway	Other
Real Time Data/Status (smart home devices)	Available	Med	-
Real Time Actuation (smart home Devices)	Available	Med	-
Access Historical data (generated by smart home devices, data aggregation/fusion)	Available		High (1)
Configuration (smart home devices)	Available	Med	-
Scheduler	Available	High (2)	-

(1)   **Data base and associated services are needed, including high capacity hard drive to keep historical data**
(2)   **Including scheduler engine running at GW level**

Table 7 Smart Home data for ADL computation

Assessment Item	Name	Smart home Contribution	Sensor Data / comments
ADL Self Performance	Bathing, personal hygiene	Yes	Presence and water consumption.
	Dressing upper body; dressing lower body	No	
	Walking, locomotion	Yes	Presence.
	Toilet transfer	Yes	Presence and water consumption.
	Toilet use	Yes	Presence and water consumption
	Bed mobility	No	
	Eating	No	

Assessment Item	Name	Smart home Contribution	Sensor Data / comments
Locomotion and Walking	Distance walked	Yes	Presence
Activity Level	Total hours of exercise or physical activity in last 3 days	Yes	Presence, mobility at home. Where has been and when
	Number of days went out of the house in last 3 days	Yes	Presence detection, arm/disarm
Activity Preferences and Involvement	Cards, games, or puzzles; crafts or arts; reading, writing, or crossword puzzles	no	
	Computer activity	yes	Smart plug, energy consumption
	Conversing or talking on the phone	no	
	Dancing	no	
	Exercise or sports	no	
	Gardening or plants?	yes	Presence, energy consumption (water)
	Helping others	no	
	Music or singing	no	
	Pets	*N/A, presence of pets is not considered in the RADIO experiment environment*	
	Spiritual or religious activities	no	
	Trips/shopping	*N/A, RADIO operates indoors*	
	Walking or wheeling outdoors	*N/A, RADIO operates indoors*	
	Watching TV or listening to radio	yes	Smart plug, energy consumption
	Discussing and reminiscing about life	no	
Time Asleep During Day	Time Asleep During Day	no	

4 WSN AND COMMUNICATION ARCHITECTURE

In this chapter, the wireless sensor network of RADIO is explained. This includes the technical requirements for the interaction between devices and gateway as well as the communication architecture within the smart home environment.

4.1 Technical Requirements on Wireless Communication

In this section we identify and analyse the main requirements based on which adequate wireless communication technologies will be adopted, configured or even extended so as to meet the objectives of the RADIO project. Respective requirements stem both from the user requirements as they are defined in WP2 as well as the purely technical demands of a state of the art wireless sensor network. It is noted that in this version we mainly focus on the main requirements axes pertinent to the first phase RADIO system prototype and that the architecture will be extended and adapted.

4.1.1 Low Power Operation

The ability of a wireless communication to promote power consumption minimization comprises a characteristic of paramount importance. Meeting such objective is a multifaceted task involving many critical aspects of a wireless communication platform concerning both hardware and software aspects. Starting with a physical medium, the respective technology must support low power transmission, receive, idle, and sleep operation modes (e.g typically transmission power in nowadays WSN technologies do not surpass 0dBm but can go as low as -30dBm). Additionally, other relative mechanisms such as modulation schemes, frequency bands, frequency channels etc. must also be selected taking into consideration power conservation.

With respect to the typical WSN software stack as depicted in Figure 11, moving up the network stack the MAC layer is also of critical importance as far as power consumption is concerned. In this context efficient, flexible and adaptive scheduling allows the network device to remain in sleep state (i.e. support of very low duty cycle) is without a doubt a cornerstone in nowadays dominant communication technologies striving towards power conservation. Of course such an objective entails highly efficient mechanisms so as to avoid robustness degradation. Additionally, optimized control operations/communication come into play also representing a substantial part of the MAC overall duties. Adequate routing support (when/if needed) can also pose important challenges with respect to power consumption, thus dominant communication technologies promote specific routing approaches further enhancing their power conservation capabilities.

Although respective standards do not indicate specific restrictions, another aspect pertaining to the system-wide low-power capabilities is related to processing units supporting any specific technology. On one hand specific architectural characteristics such as bus width, cache support, pipeline utilization, hardware accelerators support, multi-core support etc. drastically affect be trade-off between processing speed and power consumption. On the other hand, the adequate support of low power operation modes is very important in the context of the specific requirements definitions. When it comes to WSN networks and cyber-physical systems, all key companies (e.g. TI, ARM) offer solutions focusing on this aspect such as the well know TI MSP430 or/and the ARM Cortex M family (with Cortex M0 being the most power conservative representative).
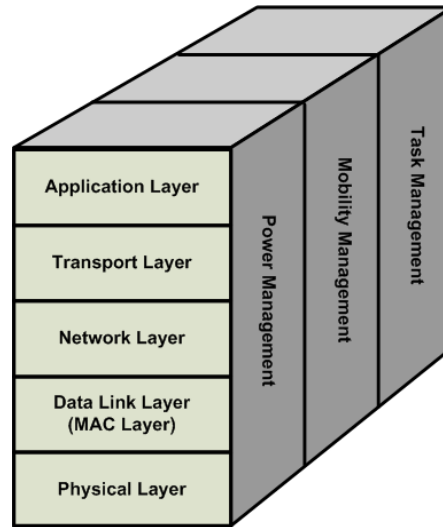
Figure 11: Wireless Sensor Network Software Stack

4.1.2   Easy and Rapid Connection Support

Typically all connections between two specific nodes are characterized as either *connection-based* or *connectionless*, with each approach offering specific advantages and disadvantages. Connectionless communication is less complex since it needs no extra mechanisms to setup/maintain/end a specific connection which also means the whole protocol is much lighter. It also claims to pose less delay by comparison to the delay a connection-based communication adds to either form a connection, or re-establish a connection. Finally, such a connectionless communication makes multi-hop and mesh networks much easier to form. Connection-less communication is typically found in IEEE 802.15.4 networks.

However, connectionless communication offers no inherent support towards robust data transfer, load balancing, QoS and in general traffic control. It is noted that such requirements are increasingly required in nowadays and future WSNs since the applications increase in complexity, criticality and communication demands. In connection based communication, the two communication nodes usually assume a specific role (e.g. master/slave) and agree on the characteristics of the communication, so as a scheduler can anticipate traffic workload and QoS demands so as to prepare an efficient packet transmission schedule. Typical communication technology traditionally following this approach is the Bluetooth technology.

Finally, it is noted that since both approaches offer important advantages in nowadays WSN application dominant communication technologies, more and more offer functionalities enabling both communication approaches. Such capabilities are expected to offer increased flexibility and must also be taken into consideration.

4.1.3   Data Throughput Capabilities

Without a doubt a data volume per specific time period a communication can handle comprises a critical metric in all network types. However, in WSN typical application scenario and specifically in RADIO the respective requirements are anticipated to be low or moderate (e.g. in the order of a few tens kbps) so as to be easily accommodated by existing solutions. Specifically the modality that will comprise the most demanding case will be low/medium quality audio which effectively will be an important guideline as to the technology and additional mechanisms that may be needed (e.g. telephony speech audio source produces 8bit samples with a frequency of 8kHz yielding a traffic of 64kbps which can be considered the upper limit of what might be required).

Of course an unpredictable factor is the concurrent traffic flows and how they are accommodated, especially the burst traffic that is typical of realistic WSN scenarios. Such scenarios will also introduce respective requirements as far as data handling is concerned.

Furthermore, throughput capabilities relate to the mechanisms a wireless technology employs to assure robust communication. By this requirement we refer to the mechanisms each communication employs so as to, on one hand, avoid packet collisions and on the other hand, accurately and rapidly identify a packet collision and recover from it. Such mechanisms include efficient clear channel assessment mechanisms, CTS/RST control packet exchange, CRC (Cyclic Redundancy Check) and support for acknowledgment mechanisms.

4.1.4 Delay - Jitter requirements

Analogous to the throughput capabilities, mean delay and respective requirements are anticipated to be relative relaxed, although this expectation remains to be confirmed by testing on use cases and scenarios specific to RADIO. Once again audio will most probably comprise the most important challenge and particularly with respect to the delay jitter that the network will exhibit. Delay jitter can prove to be important either to the understanding of a particular audio modality as well as to the processing algorithm of that audio modality by the respective algorithms demanding the packet sequence to have a specific inter-arrival delay (e.g. to accurate reproduce the actual audio).

4.1.5 Quality of Service Support

As the WSN technologies are utilized in an increasing range of applications scenarios and as the users' tend to pose increasing demands upon respective deployments, the needs for advanced quality of service capabilities is drastically increasing and is expected to do so in the future. By quality of service we refer to the differentiation among data traffic flows with respect to various (as well as dynamically changing) parameters as well as differentiation among users. With respect to such differentiation a communication technology is able to apply appropriate methodologies so as to meet the different requirements posed by each category in each particular time (to the degree that is possible). Also it is valuable to allow the developer to apply custom algorithms so as to apply specific policies. In the MAC layer, such capabilities are typically dependent upon the access scheme utilized, typically CSMA, TDMA or a combination of these two.

Carrier Sense Multiple Access (CSMA) is the most common distributed access approach where each node operates independently from each other trying to access the wireless communication medium. Being very simple to implement and not requiring any coordination was quite prominent in earlier WSN deployments were no QoS was required or/and the traffic demands were very low. However, as demands increase CSMA reveals some significant disadvantages since it is very difficult to apply algorithms to meet strict QoS demands. Furthermore, in high traffic scenarios packet loss increases exponentially leading the network congestion and even effectively communication collapse. Different versions of CSMA based approaches are quite popular among the IEEE 802.15.4 based platforms. By contrast, *Time Division Multiple Access (TDMS)* assumes a commonly known schedule which clearly defines which communication node can transmit in each particular time slot and, in general, assigns a specific role in each particular time slot of a more extended time period usually referred to as frame. As it is easily understood through such a framework, an algorithm can allocate each traffic flow appropriately so as to meeting each particular requirement and minimize the possibility of collision. Also through such mechanisms, maximum power conservation can be achieved since a node is aware of when for how long it needs to be awake. However, WSN being a typical network facing both low and possible burst traffic periods a prominent approach is to combine the two extreme approaches so as to exploit the benefits of both. Bluetooth is such an example segmenting the time in large frames and time slots when specific needs are to be met, but allowing during specific time period to all interested nodes to compete for access to the medium using CSMA approach.

4.1.6 Open Source Approach Support

As RADIO is primarily a research project, the open source approach is highly appreciated in all aspects of it. Particularly with respect to the WSN communication technology, offering accessible standards, well-defined APIs and even access to the internal functionality of respective communication protocols and mechanisms also comprise a valuable requirement. Without a doubt the degree by which aforementioned features are offered by a specific communication technology can greatly influence the configurability and extensibility of the respective solution offered in the context of RADIO. Also availability of adequate development environments play an important role, since they can help reduce the time required from the thinking a solution to actually developing it and testing in the actual environment drastically.

4.1.7 Network topology flexibility

Unpredictable, rapid and dynamic topology changes are an inherent characteristic of WSN networks. Even more, support of such conditions is an important advantage of such technologies over other typical network types. Therefore, it is a critical requirement for a respective communication technology to support or/and give the tools to develop algorithms to support such cases. On the one hand, unpredictable topology changes lead to increased/decreased of network congestion in particular areas of the network and in general changes in the communication conditions that the technology must take into consideration and adjust adequately. On the other hand, unpredictable topology supports cases were a node can't communicate directly with the intended receiver and thus intermediate nodes act as relays to forwarding data packets towards the final receiver. The communications technology must facilitate that specific nodes (or all nodes) of the netwaok can assume the role of a relay. Furthermore, the communications technology must allow for dynamically changing multi-hop data routes. This includes providing the route discovery, maintenance, and deletion functionalities of routing protocols. It is noted that in realistic WSN networks such as RADIO, a tree based data flow topology will be required so that when/if a node(s) cannot reach directly the gateway they may have the opportunity to convey respective data over multihop data paths.

4.1.8 Well accepted by Industry and Academia

Aiming towards a platform of an extended and evolving life cycle it is beneficial to rely on communication technologies that have a significant footprint in both industry and academia. Having a substantial impact and acceptance in industry through respective commercial solutions or/and being integrated as 3rd party solutions in commercial products indicates a viable and practical solution as well as adequate support can be anticipated. Also being adopted by key players comprises a critical advantages concerning longevity, extensibility as well as expandability (in different application domains, modality support etc.) of the respective communication technology. On the other hand receiving high interest from academic research groups is also important in the context of the RADIO as a research project since it indicates an open approach allowing researchers/designers/developers to further experiment and extent respective solution so as to meet specific research objectives. This leads to the formation of an extended research community which resembles a significant plus when specialized research is needed.

4.1.9 Security Support

Although left for last, security support comprises one of the most important and critical features of nowadays and future communication technologies in WSNs. This is due to the fact that WSNs' utilization is expanding to highly critical application domains such as healthcare and wellbeing of users which directly relates to safekeeping of highly private personal data as well as the adequate management and distribution of them. The RADIO project platform lays in this context since private data or/and events may be acquired, stored, wireless transmitted and processed.

Security support comprises by three services: data privacy, authentication and authorization. The first one relates to the assurance that data can be "understood" (i.e. be useful information) only by nodes that have the right to do so, while for the rest of the nodes effectively are useless bytes. This service typically relies on the capabilities of WSN nodes to execute sophisticated encryption algorithms offering high security level. Consequently inherent support of such cipher algorithms by the used communication technology is a significant added value. The second service allows nodes to prove they are who they claim to be which effectively allows accountability and is a prerequisite for the third service which effectively determines what level of access can a specific user have to certain data or/and functionalities. It is easily understood that in scenarios of RADIO which accesses personal data, habits, activates, whereabouts and environmental conditions such features can be of critical importance.

4.2 Wireless Communication Architecture

The operating environments targeted by RADIO are domestic homes of elderly people. These homes generally do not have sufficient technological infrastructure to provide ad-hoc ambient assisted living (AAL) services. In order to guarantee that the impact of the RADIO system is not limited by requiring specific communication infrastructure to be prior deployed at the end-users' homes, we consider the infrastructure connecting the RADIO components to be a substantial part of our architecture design.

Specifically, the RADIO communication architecture includes:

- The wide-area communication between each RADIO Home and remote components, such as storage and processing facilities at the hospital or notification functionalities at care-givers' devices. We do *not* delve into any further details here, as this is the subject matter of Task 5.1.
- The local, mostly (or even exclusively) wireless communication of the components deployed within each RADIO Home.

The RADIO Home environment itself comprises subgroups (Figure 13), each fulfilling a different task and designed in response to different requirements:

- Basic smart home: Off-the shelf smart home devices
- Extended smart home: Advanced devices, integrating sensing and low-energy processing, as well as the RADIO Home server
- Mobile platform: Robotic platform, integrating sensors and limited computation functionality
- User interface devices

In order to provide long term support and reliability, the *basic smart home* sensor and actuator devices will be commercial off-the-shelf components such as Z-Wave products. Z-Wave is a wireless communication protocol designed for close range sensor and actuator networks. It does not require a coordinator node as it automatically initializes a full mesh network[16]. However, in the context of RADIO coordinator nodes are used for Z-Wave in order to ensure connection robustness and reliability. This way, it is possible to modularize the architecture and expand it with additional sensors if necessary. There are sensors and actuators for data acquisition in order to react accordingly to user stimuli. However, by contrast to some current smart home architectures, the individual devices should *not* require internet connectivity for their operation. This is due to the privacy requirement that all raw data processing must be carried out within each home and only high-level event logs should be recorded at cloud storage or other remote databases. This includes not only interRAI item recognition logs, but also system operation events (such as using smart home functionalities).

The *extended smart home subgroup* will rely on Bluetooth to connect devices that carry out sensing and local processing in order to recognize daily activities and routines. Such devices cannot be

---

[16] Z-Wave Alliance. http://z-wavealliance.org/, date of access: August 2015.

developed for Z-Wave, due to its closed specifications. This creates the requirement for a gateway that interfaces between Z-Wave and Bluetooth, in order to allow RADIO to combine the robustness offered by commercial Z-Wave devices with the flexibility to develop new devices offered by Bluetooth.

Finally, the mobile platform establishes its own internal network in order to integrate its various sensing and processing elements, but also needs to connect to the overall RADIO Home. The platform is outfitted with two interfaces, the Bluetooth Low energy (BLE) and the WiFi interface. BLE connectivity is provided to support direct access to devices of the extended smart home and WiFi connectivity for data transfer requirements.

This analysis gives us the architecture in Figure 13 where the Z-Wave network, the Bluetooth network and the TCP network interface at a *central gateway* and where:

- The devices of the basic smart home connect to the Z-Wave network
- The devices of the extended smart home connect to the Bluetooth network
- The mobile platform connects to the TCP and the Bluetooth network

The gateway also serves as the central data processing point, since it is connected to all available devices in the smart home environment. Finally, the gateway interfaces with remote components of the RADIO ecosystem over the Internet.

5 OVERALL RADIO CONCEPTUAL ARCHITECTURE

In the previous chapters we have introduced the main components of RADIO, and for each one we have provided a snapshot of its current architecture and components. In this chapter, we are combining the architecture of each component to form the first version of the RADIO architecture concept.

As shown, four main components are the primary building blocks of RADIO:

1) Robot
2) Wireless sensor networks
3) Smart Home
4) ADL and mood recognition methods

The recognition methods form the main computational resource of RADIO, and it is the principal component addressing the InterRAI assessment items described in WP2.

RADIO will build upon NCSR-D's pyAudioAnalysis for feature extraction. The extracted short-term features will be fed as input to all audio-related sub-modules (audio event recognition, speech recognition, etc.)

Note that most of these are based on the following basic (pre)processing modules (Figure 12):

1. Audio feature extraction
2. Body detection
3. Face detection

In this initial version, we have estimated of the hardware requirements, both in terms of data capture (sensors; audio and video) and algorithm computation (micro and standard computers). These will be refined through testing.
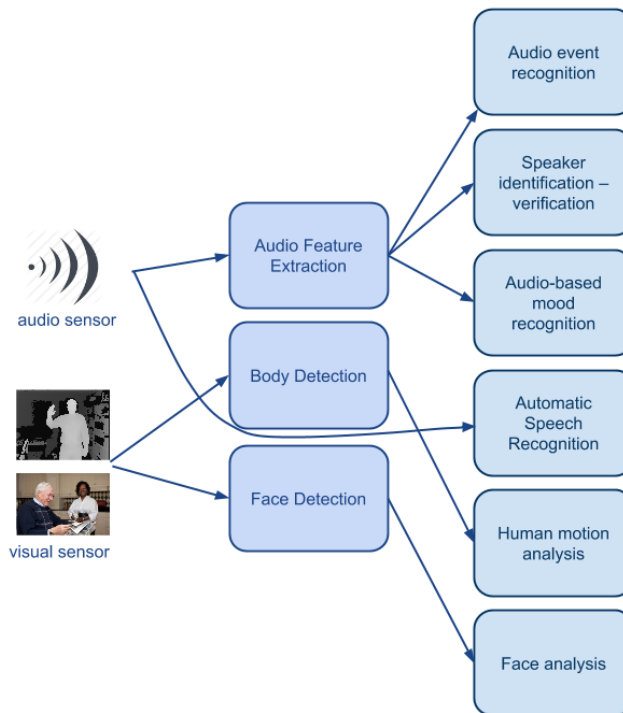

Figure 12: Conceptual architecture of the RADIO sensor data analysis components.
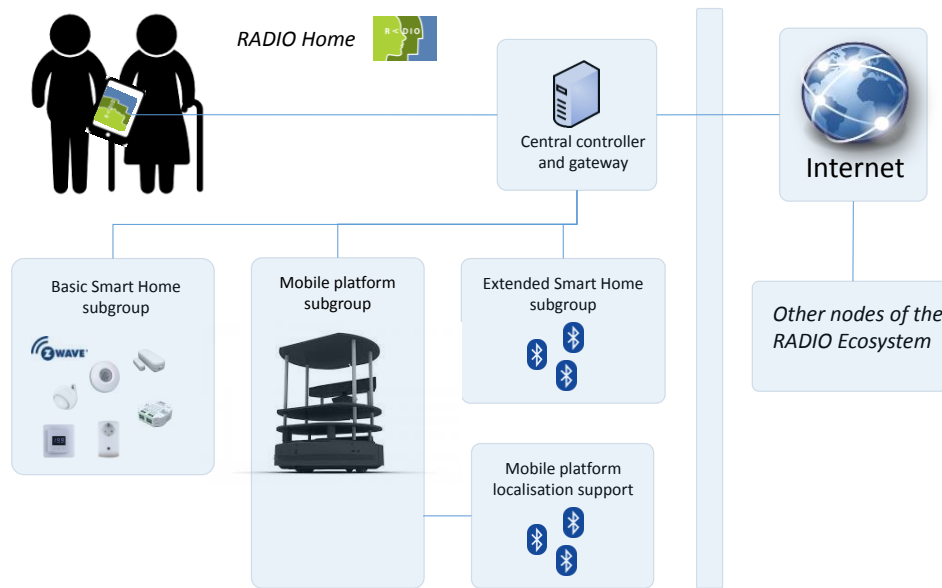
Figure 13: Conceptual architecture of RADIO Home communication channels

The Robot is an important element in the RADIO architecture concept. It is an entity that will continuously be in "motion & tracking" elder person activity. The robot is targeted to be the main interface with elderly people, but it is capable to hold computational resources and sensors needed by the recognition methods. Whether or not, the robot will hold them will be decided in future steps within WP4.

The goal of the smart home is to provide means of interaction with different controlled parts of a house, for instance open/close doors, curtains or shuts. Thus, it will provide home automation capabilities at a finger level of the elderly. The smart home thus will be a slave automation component in the RADIO architecture of the applications tailored in WP5. Furthermore, we will complement the data acquisition for the ADL and mood recognition methods, based on the huge amount of off-the-shelf sensors. Sections 2 and 3 provided an initial relation of InterRAI assessment items and RADIO Home data capture capabilities. This initial relation has been provided for the sake of future discussions, and will be updated in next iterations of this document.

All components need to be connected forming the RADIO communication system. RADIO will use common communication technologies between the main components (IP based communications, WIFI or LAN). Section 4 provided an initial overview of main important requirements for the connection of different sensors and actuators in the RADIO architecture, where low power wireless sensor technologies are the target.

Figure 13 depicts the first Conceptual architecture of the RADIO Home communication channels. Naturally, this architecture does not reflect the physical wireless communication channels which will be identified in WP4; and further extended in scope beyond the home capabilities for ADL and mood computation toward a service that will connect elderly people with doctors and caregivers that will be developed in WP5.