Advanced Information and Knowledge Processing

Hanen Idoudi
Thierry Val *Editors*

Smart Systems for E-Health

WBAN Technologies, Security and Applications

Springer

Advanced Information and Knowledge Processing

Information systems and intelligent knowledge processing are playing an increasing role in business, science and technology. Recently, advanced information systems have evolved to facilitate the co-evolution of human and information networks within communities. These advanced information systems use various paradigms including artificial intelligence, knowledge management, and neural science as well as conventional information processing paradigms.

The aim of this series is to publish books on new designs and applications of advanced information and knowledge processing paradigms in areas including but not limited to aviation, business, security, education, engineering, health, management, and science.

Books in the series should have a strong focus on information processing - preferably combined with, or extended by, new results from adjacent sciences. Proposals for research monographs, reference books, coherently integrated multi-author edited books, and handbooks will be considered for the series and each proposal will be reviewed by the Series Editors, with additional reviews from the editorial board and independent reviewers where appropriate. Titles published within the Advanced Information and Knowledge Processing Series are included in Thomson Reuters' Book Citation Index and Scopus.

More information about this series at http://www.springer.com/series/4738

Hanen Idoudi · Thierry Val
Editors

Smart Systems for E-Health

WBAN Technologies, Security and Applications

Springer

Editors
Hanen Idoudi
National School of Computer Science (ENSI)
Manouba University
Manouba, Tunisia

Thierry Val
CNRS-IRIT
University of Toulouse 2 Jean-Jaurès
Blagnac, France

ISSN 1610-3947 ISSN 2197-8441 (electronic)
Advanced Information and Knowledge Processing
ISBN 978-3-030-14938-3 ISBN 978-3-030-14939-0 (eBook)
https://doi.org/10.1007/978-3-030-14939-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Introduction

E-health has attracted a worldwide great interest since more than a decade.

E-health is an evolving concept since its inception due to the numerous technologies that can be adapted to offer new innovative and efficient e-health applications. Nowadays, with the tremendous advancement of wireless technologies, sensors and wearable devices and software technologies, new opportunities arisen and transformed the e-health field. Moreover, with the expansion of the Internet of Things (IoT), and the huge amount of data that connected e-health devices and applications are generating, it is also mandatory to address new challenges related to the data management, applications management and their security. Moreover, artificial intelligence and machine learning are becoming more and more needed in many applications to build smart, resilient and innovative medical services.

Related research areas are numerous and complementary. Work is being done on the physical layers dedicated to connected health equipment. Industrial R&D departments are increasingly working to deliver innovative products and applications, while meeting standards and taking advantage of the significant capabilities of today's wireless transmission systems. The scientific community has understood the new challenges that must now be worked on, especially to make these radio physical layers interoperable and the least harmful for patients. Investigating solutions, no longer based on electromagnetic radiation, such as light with the new versions of Li-Fi, should find in the field of e-health an interesting field of application.

Security and privacy are also an area of research with valuable and high added value for e-health. Beyond the classic and general security problems such as cybercrime, cybersecurity and cyberdefense, important research works are carried out for the very particular context of the patients' care workflow, shared health records, taking into account the individualized access of medical staff throughout life.

Finally, the use of a growing number of sensors and actuators dedicated to comfort and health already induces nowadays a very large mass of data. This trend is expected to increase in the coming years. Big data, deep learning and artificial intelligence will therefore, in all areas related to the Internet, grow strongly. In particular, IoT applications dedicated to e-health will be a major application field of these new techniques of massive data processing. Companies specializing in IT for e-health find a new area of R&D attractive, for the treatment, but also the presentation of

data from equipment dedicated to e-health, especially for the treatment of various pathologies, such as re-education.

Faced with these observations, this book has been structured into three parts representative of these three major research issues.

The Layout of this Book

The purpose of this book is to review, through ten chapters, the recent advances in e-health technologies and applications. It is intended for all practitioners (industrial and academic) in the areas of: (i) wireless communications and embedded technologies applied to e-health and (ii) software, cloud computing, artificial intelligence and big data and, (iii) security for e-health applications.

In particular, this book is folded into three parts.

In the first part, we review three recent proposals for the physical design of medical devices, signal processing and emergent wireless technologies for e-health.

The second part of the book is composed of four chapters focusing on security and privacy in IoT-based e-health applications.

The third part of this book presents three chapters reviewing applications, data mining and data analytics for e-health.

Through the ten chapters of this book, we can clearly identify that intelligent systems dedicated to e-health applications are now experiencing considerable growth.

This book is not of course exhaustive in terms of research and development actions carried out around e-health. We can also cite research works related to the identification and location of elderly people with reduced mobility, especially inside buildings, where conventional GPS-based systems or GALILEO are inoperative. Numerous smart home platforms have emerged to provide researchers, students and businesses with real-life experience. These Living Labs are also gradually replaced by the concept of True Live Lab which allows the behavioral study of users in their usual environment.

Advances in medicine have for many years been an extension of life. The resulting aging of the world's population will only increase the development needs of e-health: progress leads to progress … the loop is complete!

Hanen Idoudi
Thierry Val

Contents

Editors and Contributors

About the Editors

Hanen Idoudi is a professor in computer science at the National School of Computer Science (ENSI), The University of Manouba in Tunisia. She received her HDR (Habilitation à Diriger des Recherches) in 2017 at University of Toulouse 2—Jean Jaurès (France) and her Ph.D. degree jointly from the University of Rennes 1, France (where she was also member of IRISA, INRIA, Rennes), and from ENSI, in 2008.

Her research focuses on issues related to IoT communications protocols and wireless networking in general: Security, MAC layer optimization, routing, quality of service (QoS).

Thierry Val obtained his Ph.D. in computer science at Blaise Pascal University, Clermont-Ferrand, France, in 1993. In 1994, he became a lecturer at the University of Toulouse, where he currently teaches networks and computing systems. He obtained his HDR in 2002. He is now a full professor for the University of Toulouse 2 at the Blagnac Institute of Technology. He was submanager of the LATTIS laboratory, where he managed a research activity on wireless local networks and related protocols. He is now a member of IRIT-CNRS laboratory of Toulouse in RMESS team. His current research focuses on wireless networks in smart homes, DL-IoT (Device Layer-Internet of Things), localization of wireless connected objects. Now, he is the head of the research group of Blagnac IUT and the manager of IRIT-UT2J laboratory.

Contributors

Raluca Maria Aileni Faculty of Electronics, Telecommunication and Information Technology, Politehnica University of Bucharest, Bucharest, Romania

Monaam Ayachi Department of Electrical Engineering, University of Tunis, CEREP, ENSIT, Tunis, Tunisia

Saoussen Bel Hadj Kacem Faculty of Economics and Management of Nabeul, Carthage University, Nabeul, Tunisia;
COSMOS, National School of Computer Sciences, Manouba University, Manouba, Tunisia

Jihen Bennaceur National School of Computer Science, University of Manouba, Manouba, Tunisia

Amel Borgi Institut Supérieur d'Informatique, LIPAH, Université de Tunis-El Manar, Tunis, Tunisia

Damien Brulin LAAS-CNRS UPR8001, Université de Toulouse, CNRS, UT2J, Toulouse, France

Eric Campo LAAS-CNRS UPR8001, Université de Toulouse, CNRS, UT2J, Toulouse, France

Marie Chan LAAS-CNRS UPR8001, Université de Toulouse, CNRS, UT2J, Toulouse, France

Daniel Estève LAAS-CNRS UPR8001, Université de Toulouse, CNRS, UT2J, Toulouse, France

Aniruddha Gokhale Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN, USA

Akram Hakiri Department of Computer Science and Telecoms, University of Carthage, SYSCOM-ENIT, ISSAT, Mateur, Tunisia

Mahbouba Hattab Higher Institute of Applied Science and Technology, University of Sousse, Sousse, Tunisia

Hanen Idoudi National School of Computer Science, University of Manouba, Manouba, Tunisia

Ahmed Maalel National School of Computer Sciences, RIADI Laboratory, University of Manouba, Manouba, Tunisia;
Higher Institute of Applied Science and Technology, University of Sousse, Sousse, Tunisia

Gonçalo Marques Polytechnic of Coimbra, ESTGOH, Oliveira do Hospital, Portugal

Zouhair Mbarki Department of Electrical Engineering, CEREP, ENSIT, University of Tunis, Tunis, Tunisia

Sami Othman Faculty of Medicine of Tunis, Tunis-El Manar University, Manouba, Tunisia

Sever Pasca Faculty of Electronics, Telecommunication and Information Technology, Politehnica University of Bucharest, Bucharest, Romania

Leila Azouz Saidane National School of Computer Science, University of Manouba, Manouba, Tunisia

Arif Sari Department of Management Information Systems, School of Applied Sciences, Girne American University, Kyrenia, Cyprus

Hassene Seddik Department of Electrical Engineering, University of Tunis, CEREP, ENSIT, Tunis, Tunisia

Joshua Sopuru Department of Management Information Systems, School of Applied Sciences, Girne American University, Kyrenia, Cyprus

George Suciu Faculty of Electronics, Telecommunication and Information Technology, Politehnica University of Bucharest, Bucharest, Romania;
Beia Consult International, Bucharest, Romania

Nicolae Tapus Faculty of Automatic Control and Computers, University Politehnica of Bucharest, Bucharest, Romania

Carlos Valderrama Faculty of Engineering, Department of Electronics and Microelectronics, Mons University, Mons, Belgium

Part I
Physical Design of Medical Devices, Signal Processing and Emergent Wireless Technologies for E-Health

Chapter 1
Design and Control of Multifunctional, Multiarticulate Prosthetic Hand

Monaam Ayachi and Hassene Seddik

Abstract Limb loss can occur at different levels and for different reasons. The leading cause of limb amputation can be related to accidents like traffic accidents or work accidents; it can be also caused by disease such as diabetes, dysvascular amputation, trauma-related amputation, cancer-related amputation, and congenital-related incidences. Amputees can also be victims of military conflict and many other reasons. Loss can affect upper-limb or lower-limb at different levels. The lower-limb replacement differs from that of an upper-limb in that hand prosthesis performs wider range of movements and has more muscle tissue to be stimulated. For example, the inability to have a firm handshake may have an adverse impact on the quality of the amputee's social and vocational life. For this reason, an intelligent esthetic prosthetic hand design is important especially in a region like the Middle East in which conflict and war are increasing day after day and amputation become widely spread. Our work proposes a control system of multifunctional multiarticulate prosthetic hand based on EMG signal processing and classification. This work uses the open dataset described in Sapsanis et al. (Improving EMG based classification of basic hand movements using EMD" in 35th annual international conference of the IEEE engineering in medicine and biology society '13, 2013 [1]) to classify six basic daily hand motions acquired from five normal patients (two males and three females). This chapter is divided into four parts, the first part consists of presenting the need for this technology, problems encountered by prosthetic hand designers, specifications of existing commercialized prosthetics, and the contribution of our prosthetic hand in making this technology available and comfortable for amputees with respect to the intelligent command and esthetic look. In the next stage, we present different techniques used for pre-processing the electromyogram (EMG) signal in order to improve the classification results. Those techniques include filtering, envelope detection, and feature extraction methods in time domain, frequency domain, and time–frequency domain. Then, dimensionality reduction methods help in reducing information redundancy

M. Ayachi (✉) · H. Seddik
Department of Electrical Engineering, University of Tunis, CEREP, ENSIT, 5, Av. Taha Hussein, 1008 Tunis, Tunisia
e-mail: ayachimonaamm@gmail.com

H. Seddik
e-mail: seddikhassne@gmail.com

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_1

and increasing inter-class separability. The third section illustrates the intelligent command techniques used to classify hand movement like support vector machine (SVM), K-nearest neighbours (KNN), artificial neural network (ANN), and linear discriminant analysis (LDA). The results are promising and some validation techniques are used to verify the consistency and the reliability of these results. The final part is the implementation of our prosthetic hand. In this stage, we describe technical specifications of this prosthesis and the integration of different parts of the whole system, software, and hardware. We discuss also the relevant feature that makes this prosthetic one of the most appreciated design.

1.1 Introduction

Nowadays, biological signals are widely used by many applications like heart attack prediction or anomaly prediction, controlling exoskeleton, prosthetic hand [2], robotic arm, drones, and many other objects.

Biological signals [3] refer to the electrical activity of human body organs, as example we can site electrocardiogram (ECG) generated by the heart, electrogastrogram (EGG) generated by stomach muscle, electroencephalogram (EEG) generated by the brain cells (neurons), electrooculogram (EOG) generated by the cornea-retinal of the human eye, and the electromyogram (EMG) generated by body skeletal muscles. In this chapter, we are interested in EMG signal as a famous method for controlling upper-limb prosthesis. Some methods use EEG for controlling prosthesis, but the use of brain control interface (BCI) is not a practical solution due to the size of the installed system.

EMG signals are generated by skeletal muscles during neuromuscular activities and detected using two principal method, and each has some drawbacks.

The first method is non-invasive in which surface electrodes are placed on the skin surface of the patient where the muscle activity is important. This is the most used method for recording EMG signals. In spite of its viability and the fact that it does not require surgery, some phenomena make the acquisition task hard, like crosstalk which refers to the phenomenon that occurs when the EMG signal from one muscle interferes with that of another limiting the reliability of the signal.

The second technique is an invasive method. Needle electrodes are implanted into the patient's body. The targeted muscle re-innervation is responsible for the production of the specific hand gesture which is picked by an electrode that eliminates the crosstalk problem. This method, however, needs surgical operation.

Due to his availability and ease of use, surface electrodes are considered to be the best method for recording and analyzing EMG signals. Previous work demonstrates that optimal identification accuracy can be achieved with two to four acquisition channels.

In this chapter we use an open dataset containing six hand gesture (cylindrical, hook, tip, spherical, palmer, and lateral) (Fig. 1.1).

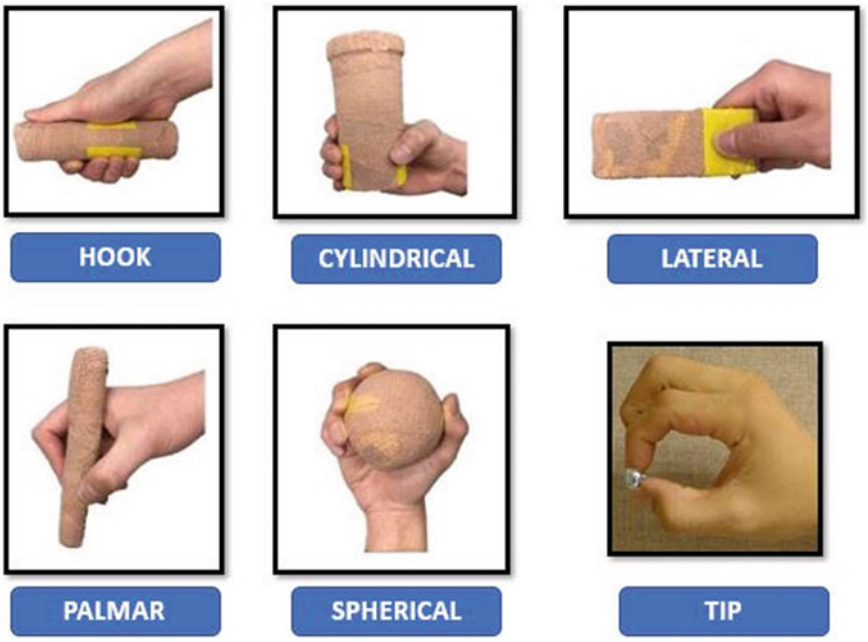


Fig. 1.1 Six studied hand gestures

Signals are collected using two forearm surface electrodes held under the flexor capri radialis and the extensor capri radialis, longus, and brevis. Other clinical experiments use four electrode collecting signal from extensor digitorum, extensor carpi radialis, palmaris longus, and flexor carpi ulnaris which are responsible for hand motion.

As described in [1], data were collected from six healthy subjects (two males and three females) approximately of the same age (20–22 years old), and every subject was asked to perform the six movement for six seconds, 30 time with two channels.

Results of the classification of those signals are high and stable. The accuracy reaches 98.05% with linear discriminant analysis, 98.3% with support vector machine, 94.17% with artificial neural networks, and 96.94% with K-nearest neighbour. The classification is performed on the envelope of the signal which is detected by four methods which are Hilbert envelope, peak-based envelope, RMS based envelope, and linear envelope.

1.2 EMG Signal Pre-processing and Envelope Detection

Pre-processing is considered to be the most important stage in the process of EMG identification and classification. Due to the low amplitude of EMG signal (between 0 and 10 mV) and the technology used to acquire it, many noise contaminations can

be harmful to the information contained within. To solve this issue, many filtering techniques are proposed in literature. In this section, we will present those type of noises and the most effective way to prevent the loss of EMG information.

1.2.1 Different Category of EMG Noise

EMG signal are sensible to many environmental changes [4] like

1.2.1.1 Electrical Noise from Power Lines and External Sources

This noise is generated by line interference from power lines (50 or 60 Hz) which is hard to remove if it is a low frequency [5], inherent noise from electromagnetic field of electronic equipment, and fluorescent light. This can be avoided by using high-quality equipment or some software solutions. Figures below illustrates EMG signal with (Fig. 1.2) and without (Fig. 1.3) electrical noise from power lines.

1.2.1.2 Motion Artifacts

Motion artifact [6] are caused by the relative movement of the electrodes with respect to the skin under target muscle. The frequency of this type of noise is usually between 1 and 10 Hz.

Motion artifact can result from a direct impact to the sensor or the body, a rapid movement of the body segment to which the sensor is attached and/or alterations in the chemical balance of the skin–electrode interface due to volumetric changes during shortening and stretching during muscle contraction. Figures below illustrate the impact of motion artifact on EMG signal (Fig. 1.4).

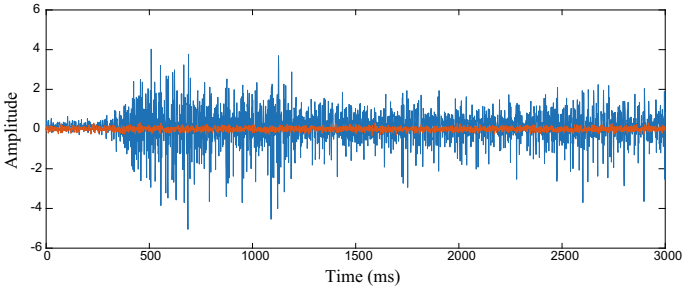


Fig. 1.2 EMG signal contaminated by line interference (shown in red)

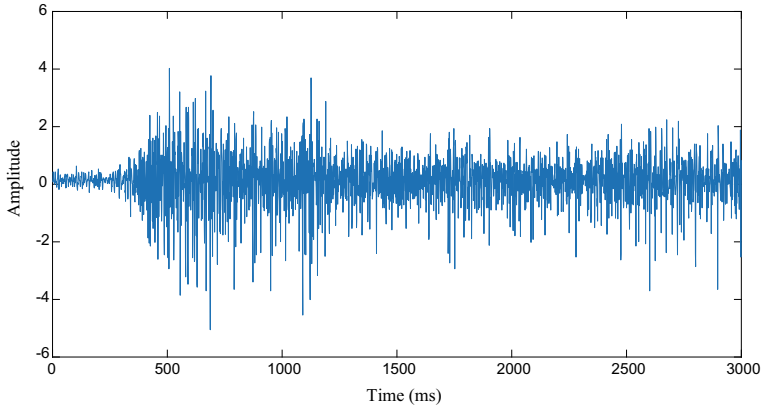


Fig. 1.3 EMG signal without line interference

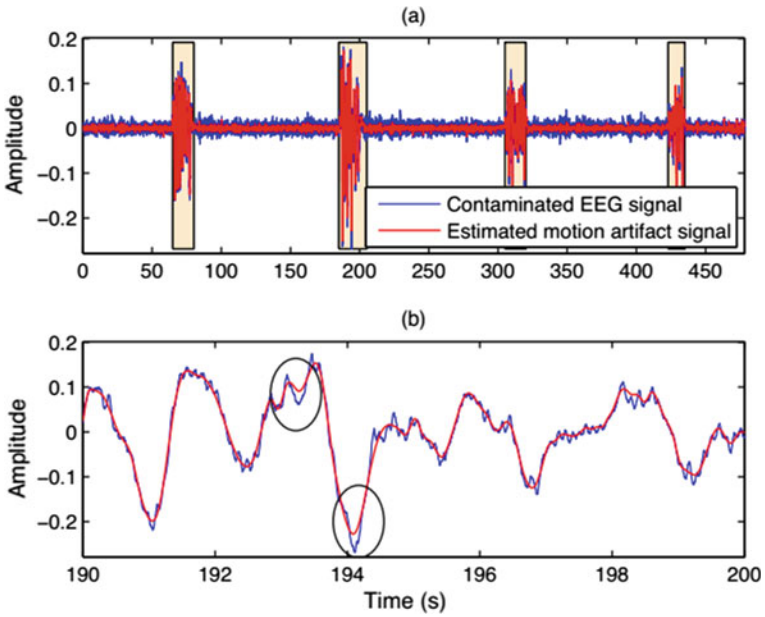


Fig. 1.4 EMG signal contaminated by motion artifact [6]

1.2.1.3 Crosstalk Contamination

Crosstalk refers to the phenomenon that occurs when the EMG signal from one muscle interferes with that of another. This contamination can be avoided by optimally choosing the electrode size and the inter-electrode distance. [7] Figures below illustrate the crosstalk contamination of EMG signal (Figs. 1.5 and 1.6).

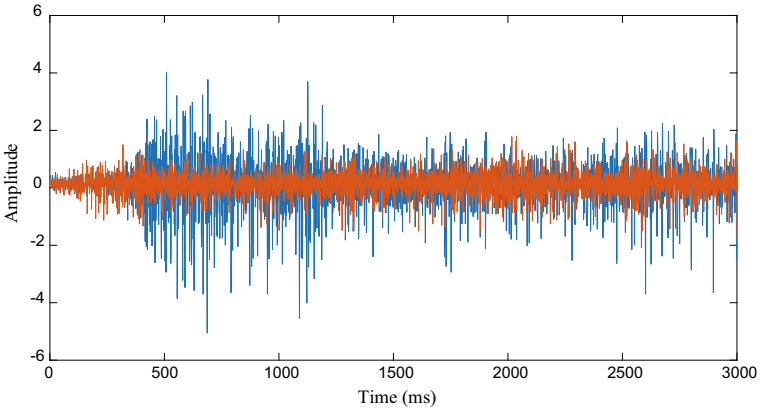


Fig. 1.5 EMG signal contaminated by crosstalk signal (shown in red)

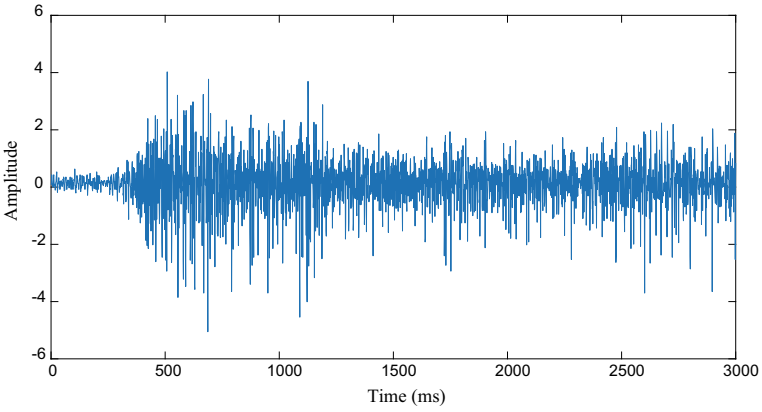


Fig. 1.6 EMG signal with no crosstalk contamination

1.2.1.4 Clipping

This term [8] describes the phenomenon of electrode saturation due to the excessive EMG amplitude as illustrated in Fig. 1.8 below (the original signal is given in Fig. 1.7). This can be avoided by reducing the amplification if possible, or reducing the EMG amplitude by changing the location of the electrode on the muscle.

Fig. 1.7 EMG signal showing regions where there is an amplitude clipping

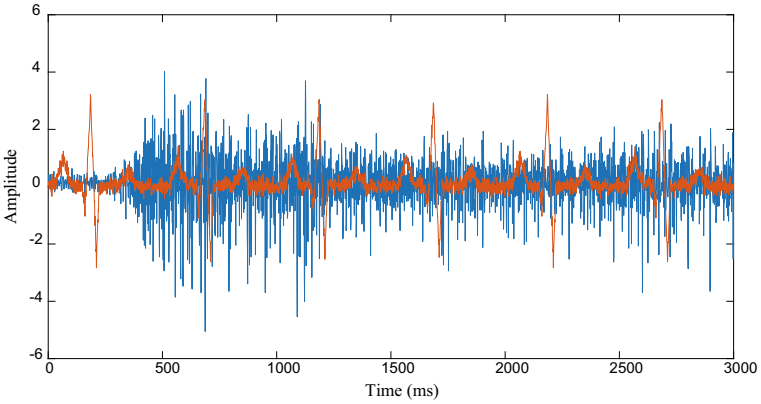


Fig. 1.8 EMG signal contaminated by a noisy heart rate

1.2.1.5 Physiological Noise

Physiological signals are all other types of signal generated by the human body except muscle. EMG are specifically contaminated by EKG signal which refers to the electrical signal generated by heart. This noise can be eliminated by optimally choosing the right place for electrodes, and Fig. 1.9 illustrates this phenomenon (the original signal is given in Fig. 1.7).

1.2.2 Denoising

With such a noisy signal, filtering is very important. Many technics can be used like low-pass differential filtering, adaptative noise cancelation, wavelet denoising, and

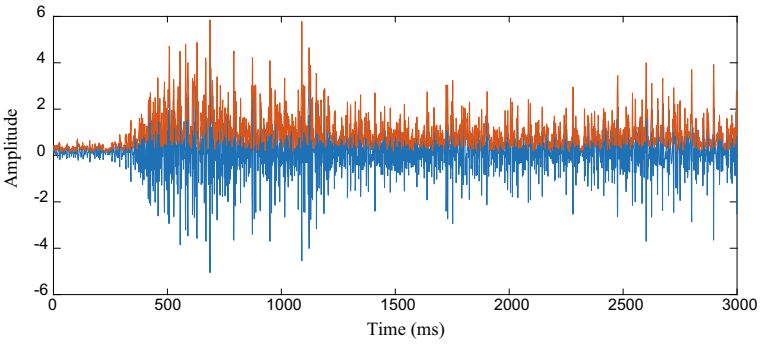


Fig. 1.9 Hilbert envelope detection (envelope red, original signal in blue)

many other intelligent tested technics like using neural network [9] and [10] and PCA or ICA [11] to reduce the impact of this noise on the original signal.

For many applications, Butterworth [12] and [13] filter prove itself useful due to it simple algorithmic implementation and low calculation effort. For this reason, for denoising purposes, we used Butterworth filter with same configuration used in [14] (a low cut-off frequency of 15 Hz and high cut-off frequency of 500 Hz) combined with a notch filter at 50 Hz to remove line interference from power lines.

1.2.3 Envelope Detection

1.2.3.1 Hilbert Transform

Hilbert transform presents many advantages by preserving signal energy and giving a relationship between real and imaginary part of a signal. Despite this, it can be computationally heavy and need additional pre-processing steps to remove noise.

The Hilbert transform [15] of a signal $x(t)$ is defined by:

$$h(t) = \mathrm{H}(x(t)) = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau = x(t) * \frac{1}{\pi * t} \tag{1.1}$$

The envelope of the signal is determined by calculating the magnitude of the complex analytical signal.

$$Z(t) = x(t) + jh(t) \tag{1.2}$$

Then, the envelope is obtained by calculating the module of $Z(t)$:

$$E(t) = |Z(t)| = \sqrt{x(t)^2 + h(t)^2} \tag{1.3}$$

Hilbert uses a four-step algorithm (MATLAB):

- Calculate the FFT of the input sequence, storing the result in a vector x.
- Create a vector h whose elements $h(i)$ have the values:
- 1 for $i = 1, (n/2) + 1$.
- 2 for $i = 2, 3, \ldots, (n/2)$.
- 0 for $i = (n/2) + 2, \ldots, n$
- Calculate the element-wise product of x and h.
- Calculate the inverse FFT of the sequence obtained in step 3 and return the first n elements of the result.

Figure below demonstrate the Hilbert envelope (red) of an EMG signal corresponding to a cylindrical gesture (blue).

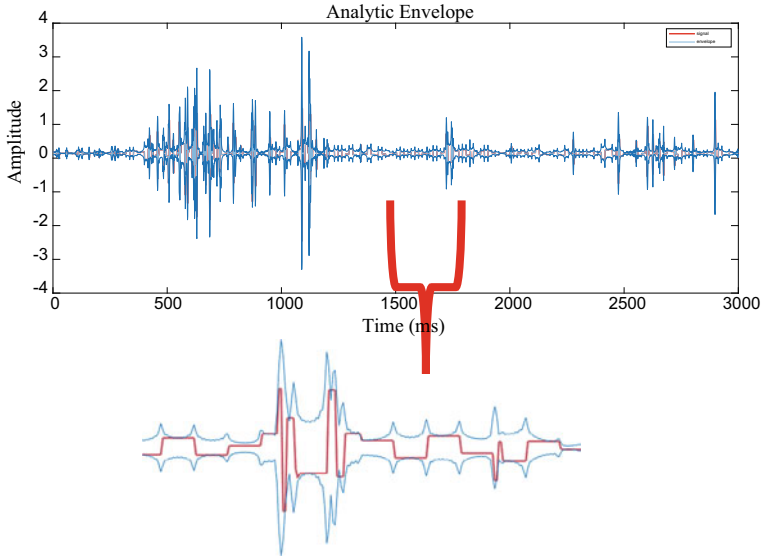


Fig. 1.10 Hilbert envelope after wavelet denoising (The red plot represents the wavelet, whereas the blue represents the envelope)

The Hilbert envelope is computationally heavy and does not detect a smooth envelope for a signal like EMG.

After denoising the signal using wavelet Daubechies with level 4, the envelope becomes clearer but still presenting fast slope variation. Zooming on a part of the signal illustrates the slope changes (Fig. 1.10).

1.2.3.2 Peak Detection

The envelope is detected by an interpolation of local maxima of the signal.

The envelope detected with peaks is more accurate than the Hilbert envelope but still not perfectly smooth. Figure below illustrates the fact (Fig. 1.11).

1.2.3.3 RMS Calculation

The envelope is detected by calculating the root mean square (RMS) within a sliding window.

The root mean square of a vector x reflects the mean power of a signal and given by the formula:

$$\mathrm{RMS}(x) = \sqrt{\frac{\sum_{n=1}^{N} |x_n|^2}{N}} \tag{1.4}$$

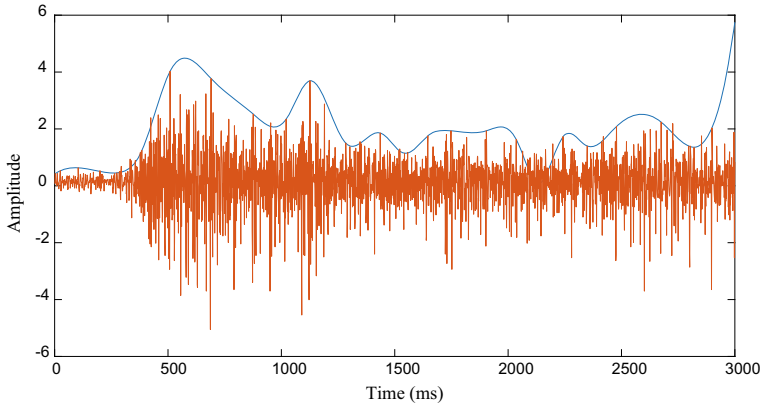


Fig. 1.11 Envelope signal based on peak detection

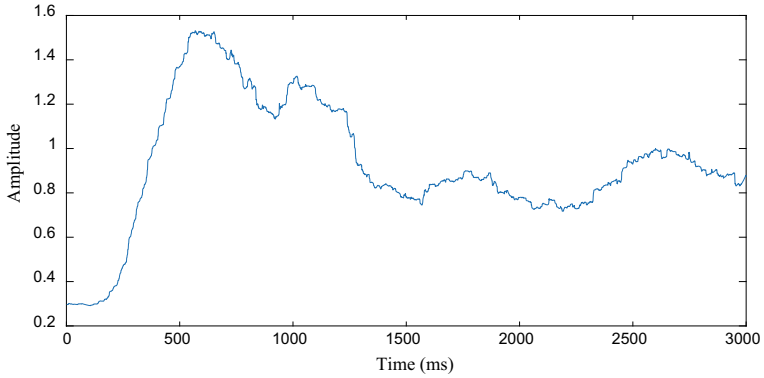


Fig. 1.12 Envelope detection based on RMS calculation

The same thing for RMS envelope, the envelope is better than both Hilbert and peak methods, but still not smooth (Fig. 1.12).

1.2.3.4 Linear Envelope

A linear envelope is obtained by rectifying the signal then low-pass filtering. A full-wave rectification means taking the absolute value of the signal which gives its shapes. Then, low-pass filtering the signal by Butterworth or Chebyshev or a sliding window with mean calculation generates the envelope.

This technique gives the best envelope compared to previous techniques (Fig. 1.13).

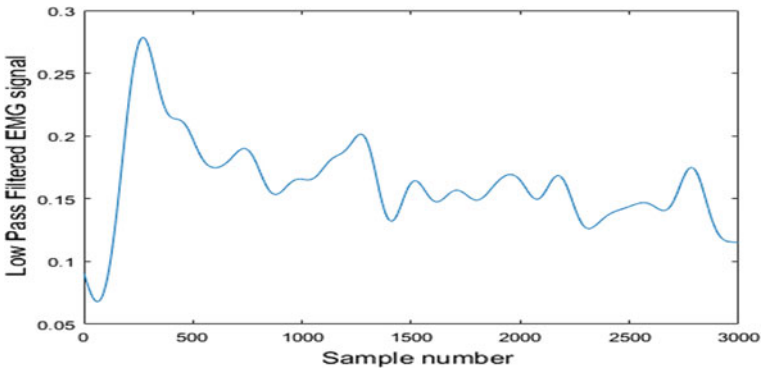


Fig. 1.13 Linear envelope detection

1.2.4 Feature Extraction

Feature extraction is the application of some formula and techniques aiming the extraction of relevant information from the signal and attenuating the effect of noise in the classification decision which makes the classifier stable and effective with future data classification. Some dimensionality reduction techniques can also be used to eliminate redundancy in data or projecting data in a lower or higher space dimension to increase inter-class separability. Three categories of feature extraction exist: the temporal domain [16], the frequential domain, and the time–frequency domain [17].

1.2.4.1 Time Domain

Integrated EMG (IEMG):

It is the summation of absolute values of a signal sample usually used to determine the on-set of EMG signal.

$$\text{IEMG}_K = \sum_{i=1}^{N} |x_i| \tag{1.5}$$

where x_i refers all values of segment k of a length N.

Mean Absolute Value (MAV) [18]:

It is the mean of IEMG also defined as the area under EMG signal after making all its values positive or rectified. Given by:

$$\mathrm{MAV}_K \frac{\mathrm{IEMG}_K}{N} = \frac{\sum_{i=0}^{N} |x_i|}{N} \tag{1.6}$$

Modified Mean Absolute Value 1 (MMAV1):

It is a linear-weighted MAV.

$$\begin{aligned} \mathrm{MMAV1}_k &= \frac{1}{N} \sum_{i=1}^{N} w_i |x_i| \\ w(i) &= \begin{cases} 1, & 0.25N \leq i \leq 0.75N \\ 0.5, & \text{otherwise} \end{cases} \end{aligned} \tag{1.7}$$

Modified Mean Absolute Value 2 (MMAV2):

It is a nonlinear-weighted MAV.

$$\begin{aligned} \mathrm{MMAV2}_k &= \frac{1}{N} \sum_{i=1}^{N} w_i |x_i| \\ w(i) &= \begin{cases} 1, & 0.25N \leq i \leq 0.75N \\ \frac{4i}{N}, & 0.25N > i \\ \frac{4(i-N)}{N}, & 0.75N < i \end{cases} \end{aligned} \tag{1.8}$$

Mean Absolute Value Slope (MAVS):

It is defined as the difference between MAV_{k+1} and MAV_k.

$$\mathrm{MAVS}_k = \mathrm{MAV}_k + 1 - \mathrm{MAV}_k \tag{1.9}$$

Root Mean Square (RMS) [19]:

It is used to estimate the power of a contraction in a non-fatiguing situation.

$$\mathrm{RMS}_k = \sqrt{\frac{\sum_{i=1}^{N} |x_i|^2}{N}} \tag{1.10}$$

Variance (VAR) [19]:

It is used to estimate the deviation of a series of EMG values from the mean, given by:

$$\mathrm{VAR}_k = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 \tag{1.11}$$

$\overline{x}$ is the mean of segment k.

Wave form Length (WL) [19] and [20]:

It is the measure of the length of a waveform, given by:

$$WL_k = \sum_{i=1}^{N-1}|x_{i+1} - x_i| \tag{1.12}$$

Zero Crossings (ZC) [21]:

It is the number of times that the signal change sign by crossing the zero. A threshold ε can be introduced to eliminate noise effect.

$$\{x_i > 0 \text{ and } x_{i+1} < 0\} \text{or } \{x_i < 0 \text{ and } x_{i+1} > 0\} \text{ and } |x_i - x_{i+1}| \geq \varepsilon \tag{1.13}$$

Slope Sign Changes (SSC) [22]:

An incremented variable used to save the number of times the slope of EMG signal changes, given by:

$$\begin{gathered}\{x_i > x_{i-1} \text{ and } x_i > x_{i+1}\} \text{or } \{x_i < x_{i-1} \text{ and } x_i < x_{i+1}\} \\ \text{and} \\ |x_i - x_{i+1}| \geq \varepsilon \text{ or } |x_i - x_{i-1}| \geq \varepsilon\end{gathered} \tag{1.14}$$

Willison Amplitude (WAMP) [23]:

This feature gives an indication of the muscle contraction level. Mathematically, it is the number of times that the difference between two consecutive values becomes greater than a given threshold.

$$\begin{aligned} WAMP_K &= \sum_{i=1}^{N-1} f(|x_i - x_{i+1}|) \\ f(x) &= \begin{cases} 1, & x > \varepsilon \\ 0, & \text{otherwise} \end{cases} \end{aligned} \tag{1.15}$$

Simple Square Integral (SSI) [22]:

It is given by the sum of squared values of EMG signal.

$$\mathrm{SSI}_k = \sum_{i=1}^{N} \left(x_i^2\right) \tag{1.16}$$

Auto-Regressive Coefficient (AR) [24]:

Auto-regressive is a time-series model in which samples are estimated using a linear combination of older ones.

$$x_k = -\sum_{i=1}^{N} a_i x_{k-1} + e_k \tag{1.17}$$

1.2.5 Frequency Domain

Frequency Median (FMD) [25]:

The median frequency is defined as frequency that divides the power density spectrum in two regions having the same amount of power. The frequency median is given by:

$$F_{MD} = \frac{1}{2} \sum_{i=1}^{M} \mathrm{PSD}_i \tag{1.18}$$

where M is the length of the power spectrum density and PSD_i is the ith power spectrum density.

Frequency Mean (FMN) [25]:

The mean frequency is that frequency where the product of the frequency value and the amplitude of the spectrum is equal to the average of all such products throughout the complete spectrum. The frequency mean is given by:

$$F_{MN} = \frac{\sum_{i=1}^{M} f_i \mathrm{PSD}_i}{\sum_{i=1}^{M} \mathrm{PSD}_i} \tag{1.19}$$

where M is the length of the power spectrum density, $f_i = (i * \text{sampling}_{\text{rate}})/(2 * M)$, and PSD_i is the ith line of the power spectrum density.

Modified Frequency Median (MFMD) [25]:

It is defined as the frequency at which the spectrum is divided into two regions with equal amplitude. The modified median frequency is given by:

$$\text{MFMD} = \frac{1}{2}\sum_{i=1}^{M} A_i \tag{1.20}$$

A_i is the amplitude spectrum at frequency bin *i*.

Modified Frequency Mean (MFMN) [25]:

It is defined as the sum of the products of the frequency f_i and the amplitude spectrum A_i.

$$\text{MFMN} = \frac{\sum_{i=1}^{M} f_i A_i}{\sum_{i=1}^{M} A_i} \tag{1.21}$$

f_i is the frequency of the spectrum at frequency bin *i*.

1.2.6 Time–frequency Domain

Short Time Fourier Transform (STFT) [22]:

STFT is used for time–frequency analysis for spectral display purposes.

$$\text{STFT}_x(t, w) = \int W^*(\tau - t)x(\tau)e^{-jw\tau}d\tau \tag{1.22}$$

Wavelet Transform (WT) [22]:

Wavelets are computationally heavy for complex signal like EMG but can be used for denoising or for EMG signal decomposition. The signal is scaled and shifted with a mother wavelet function.

$$W_x(a, b) = \int x(t)\left(\frac{1}{\sqrt{a}}\right)\psi^*\left(\frac{t-b}{a}\right)dt \tag{1.23}$$

$x(t)$ is the input signal.
ψ^* is the complex conjugate of the mother wavelet function.
a is the scale factor, and b is the shift time.
$\psi^*\left(\frac{t-b}{a}\right)$ is shifted by time b and scaled by factor a wavelet.

Wavelet Packet Transform (WPT) [22]:

Wavelet packet transform is considered as an improvement for the wavelet transform. It can improve feature space visualization, decrease noise interference, reduction of signal dimension, and used also for signal reconstruction.

1.3 Intelligent Control

Four types of intelligent learning techniques are actually used for classification and regression purposes. We can find supervised algorithms like support vector machine (SVM), K-nearest neighbour (KNN), artificial neural networks (ANN), linear discriminant analysis (LDA), and others [26, 27], in which a labeled set of data are provided to train the classifier, and its only role is to determine a model that better fits the given data and capable of correctly classifying new future information. Unsupervised learning is used when there is a large amount of data, and labeling it is hard to do, so in this kind of situation, an unsupervised algorithm is supposed to find groups of data that have some similarity and cluster them. Some unsupervised algorithms that can be used are *k*-means, hidden Markov model (HMM), Gaussian mixture model (GMM), and others. Semi-supervised learning is a learning technique in which datasets are not all labeled, and unlabeled data can carry important information [28]. Semi-supervised learning can learn from labeled and unlabeled data. The last category of learning is reinforcement learning which is based on external feedback, in other words when interacting with environment, the algorithms receive a positive and negative feedback, and based on those feedbacks the algorithm learn [29].

1.3.1 Support Vector Machine (SVM) [30]

Support vector machine is powerful supervised discriminative classifier which uses hyperplanes which are decision boundaries to separate classes. Given a set of labeled data, the SVM tries to find the optimal hyperplane that separates the data. The optimal hyperplane is determined using margins which are built based on support vectors, which refers to data points close to the hyperplane, to calculate distance between dataset and hyperplanes. By maximizing margins, we obtain hyperplanes that best separate classes. So, we ensure that future data can be classified correctly with some confidence. Figure below gives an example of good and bad margin (Fig. 1.14).

Sometimes, the separation is linearly impossible, so SVM can use kernels to pass to a high-dimensional space in which classification is better. The figure below describes the process of transformation from a low-dimensional space to a high-dimensional space (Fig. 1.15).

In the origin form, SVM is binary classifier that can classify only two classes of data by assigning 1 or −1 value to each class according to:

$$w \cdot x_i + b + \xi_i \geq 1, \text{ if } y_i = 1 \tag{1.24}$$

$$w \cdot x_i + b + \xi_i \leq -1, \text{ if } y_i = -1 \tag{1.25}$$

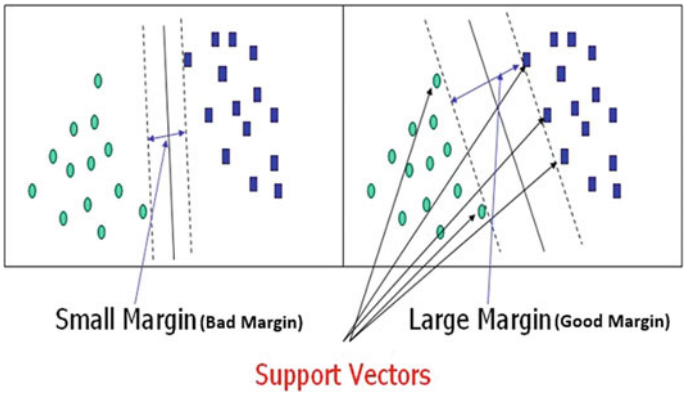


Fig. 1.14 Illustration of the difference between a good and a bad margin in support vector machine

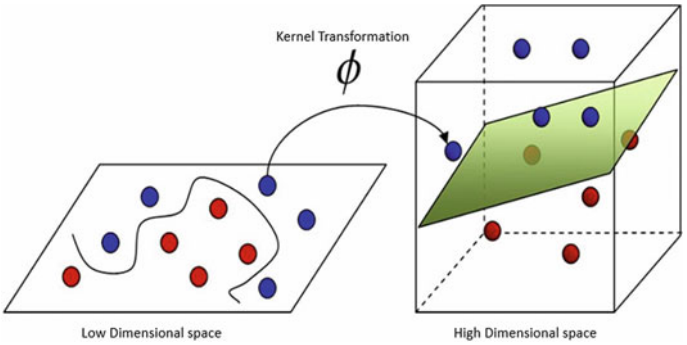


Fig. 1.15 Illustration of kernel transformation from a low-dimensional space to a high-dimensional space

w is the normal to the chosen plane.
b is the intercept term which maximize the margin.
ξ is the distance of the misclassified points from hyperplane.

With some techniques, it can easily pass to multi-class classification like one-vs-all, one-vs-all and error correction codes (ECOC). One-vs-one method is used in this chapter to transform the binary SVM classifier to a multi-class classifier with the determination of the best regularization parameter and gamma parameter.

For a classification problem with n classes, the one-versus-one strategy will build $n^*(n-1)/2$ binary classifier. The decision taken after this is by using a majority vote which assign a label to the new data.

1.3.2 *Artificial Neural Network (ANN) [30] and [31]*

Artificial neural network is invented to simulate biological neurons of the brain. A biological neuron is an electrical excitable nerve cell which receive information from other neurons via dendrite, process them in the cell body (soma), and send them to other neurons via axon terminal. The figure below describes the structure of a biological neuron (Fig. 1.16).

An artificial neural network is a network of an information processing units called artificial neurons. It is an arithmetic summation of weighted inputs ($w_i.x_i$) passed to an activation function. It can be described as above (Fig. 1.17).

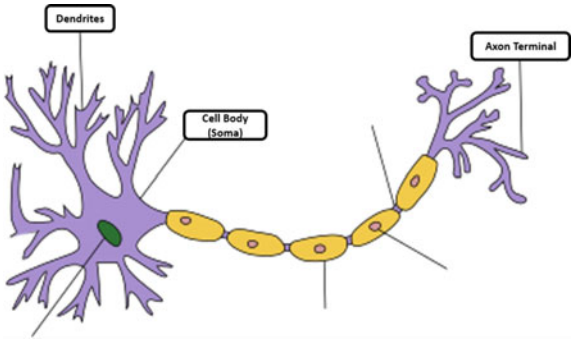


Fig. 1.16 Structure of a biological neuron

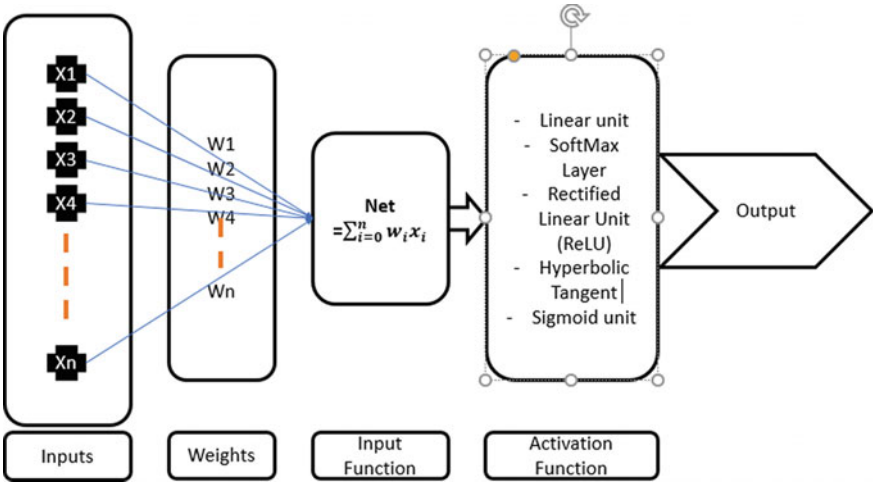


Fig. 1.17 Architecture of an artificial neural network

Table 1.1 Different activation functions

Linear unit	$f(x) = w \cdot x + b$
Sigmoid unit	$f(x) = \frac{1}{1+e^{-x}}$
SoftMax layer	$f(x) = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x_i}}$
Rectified linear unit (ReLU)	$f(x) = \max(0, x)$
Hyperbolic tangent	$f(x) = \tanh(x)$

$$\text{net} = \sum_{i=0}^{n} w_i x_i$$

Many activation functions can be used depending on the learning problem to enhance the performance of the neural network like linear unit, sigmoid unit, SoftMax layer, rectified linear unit (ReLU), and hyperbolic tangent.

The table below describes mathematic formulation of different activation functions (Table 1.1).

1.3.3 K-Nearest Neighbour (KNN) [24]

K-nearest neighbour is the most simple and basic supervised classifier. Using a labeled data, the model generated by KNN uses a majority vote technique to classify new data. In other words, the new points of data will be associated to the class that apparats the most in its nearest. KNN is based on the calculation of different type of distance between data points, so it can be a real time consuming. It is recommended that KNN should be used with a little dataset. The figure below explains that how the KNN works (Fig. 1.18).

The new data in black will be associated to the class that apparats the most in its nearest, which are the green data.

1.3.4 Linear Discriminant Analysis (LDA) [23, 31]

Linear discriminant analysis is a supervised multi-class classification algorithm that can be also used as a dimensionality reduction technique helping in pre-processing data to avoid overfitting.

Linear discriminant analysis is based on the estimation of the means and covariance matrix of input variables. The mathematic formulation of the linear discriminant analysis is given below:

$$P(y = k|X) = \frac{P(X|y = k)P(y = k)}{P(X)} = \frac{P(X|y = k)P(y = k)}{\sum_l P(X|y = l)P(y = l)} \tag{1.26}$$

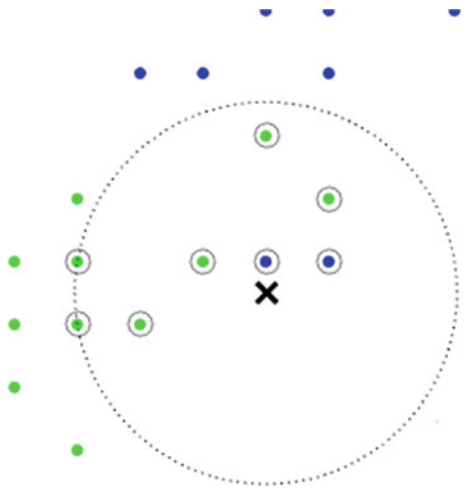


Fig. 1.18 Illustration of K-nearest neighbour classification technique

where $P(X|y = k)$ is the class conditional distribution obtained of the data for each class k selected in a manner that maximizes this conditional probability.

Prediction is then obtained by applying Bayes rule.

1.4 Software and Hardware Implementation

In the last few years, the development of prosthetic hand passed from body empowered prosthesis to myoelectrical prosthesis. The number of degrees of freedom was limited to one degree, but with the technology evolution and minimization of electronic and mechanical component size, the development of new prosthetic hands with more degrees of freedom become possible.

Many prosthetic hand manufacturers are actually emerging in the commercial market. Many products have been developed like BeBionic hand of RSL Steeper, i-LIMB hand of Touch Bionics, and Michelangelo hand of OttoBock. Those revolutionary prosthetic hand are able to perform more than opening or closing movement. They use intelligent command strategies with the conservation of the esthetic aspect to mimic the most of the human hand. The number of degrees of freedom is still limited compared to the biological hand (20 degrees of freedom), but research is becoming more intensive in this field, and many new products are emerging to the market.

1.4.1 Mechanical Design

The mechanical design should be esthetic and flexible in order to make the user comfortable. The prosthetic hand will allow multiple grasp configuration and give the possibility for introducing new hand gesture in future development.

SolidWorks 2016 × 64 Edition SP2 software of Dassault Systems is chosen to be the tool used for designing the hand (Fig. 1.19) because of the rich functionality that it contains. Basically, surfacing tools are used to have more control on the complex shape and look of the prosthetic hand. Then, design is converted to blocs ready to be 3D printed and tested.

Before talking about the prosthesis, we will describe the biological human skeletal hand anatomy [32] to have an idea about part of our prosthetic device.

The human hand is the greatest tool used for building civilizations. It had five fingers, a thumb finger, index finger, middle finger, ring finger, and little finger. It contains three principle parts: phalanges, metacarpal bones, and carpal bones. Those parts are covered by tissues of muscles and ligament which gives the hand its stability, precision, and strength (Fig. 1.20).

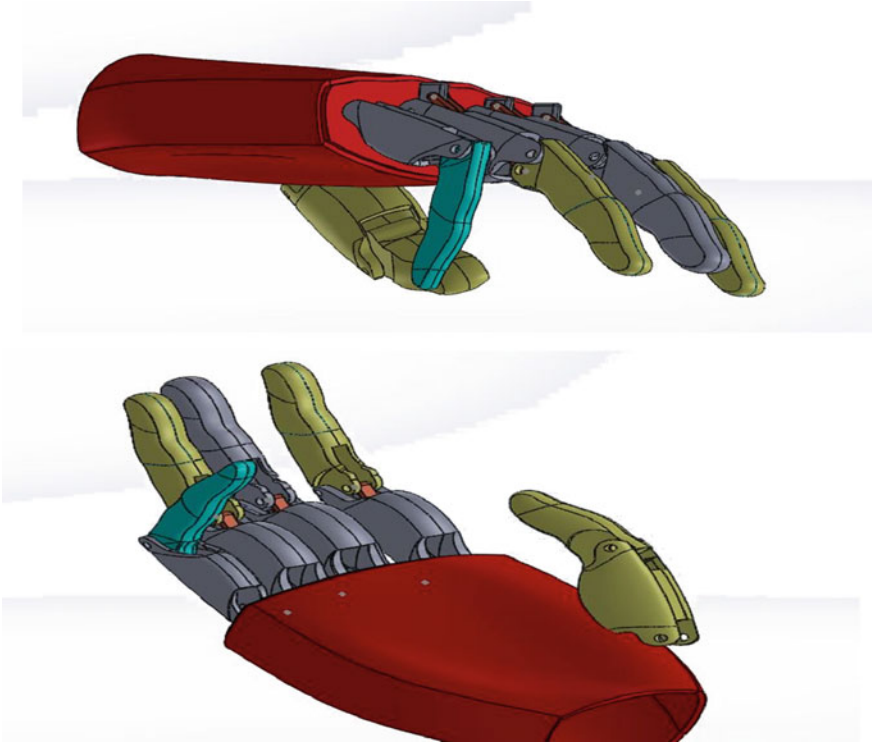


Fig. 1.19 Prosthetic hand design in solidworks

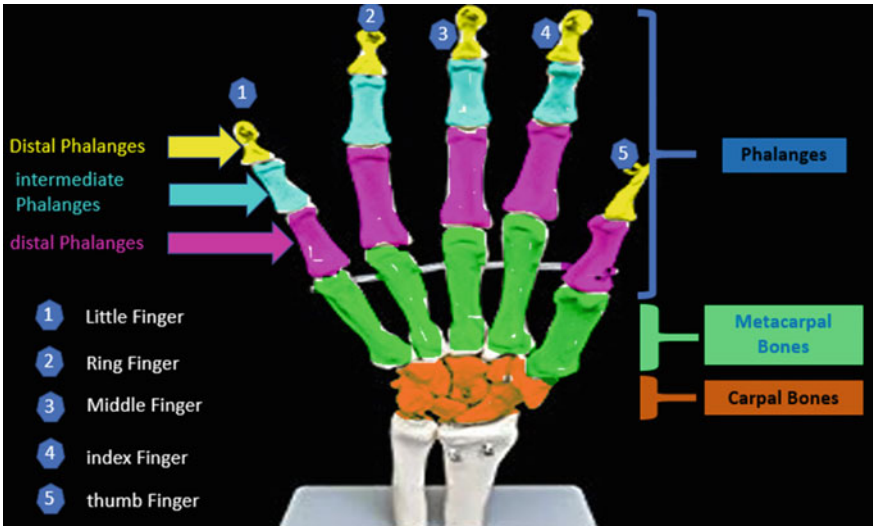


Fig. 1.20 Anatomy of the biological human hand

The carpal and metacarpal bones will not be detailed because those two parts are combined in the prosthesis design for implementing motors and related controls. The phalanges parts are not identical for all fingers. Only the thumb finger contains two phalanx, distal, and proximal; and the rest of the fingers contains three phalanx which are, from the extremity, distal, intermediate, and proximal.

Distal and intermediate phalanx are combined in the prosthetic hand to make them resistant. The movement of the distal and intermediate parts relative to the proximal phalanx is achieved by a four-bar linkage used in many real systems. Many achievable grasps can be sited, like power grasp, precision grasp, hook grasp, lateral grasp, tip grasp, and finger-point grasp. Kinematic model of finger joint coupling mechanism is given by figure below (Fig. 1.21).

Gear arrangement are very important in a system with space constraint. The proximal phalanx uses a worm gear for flexing. The principal advantage of using such a gear arrangement, is that it can block the movement of the finger at any position it takes (Fig. 1.22).

1.4.2 Electrical Requirement and Specification

In order to control the above-cited electro-mechanical conception of the hand, we are constraint to equip it with electrical devices able to manage it. The actuators have to be controlled in real time with high precision and accuracy. This is on which we are working actually in order to bring this project to life.

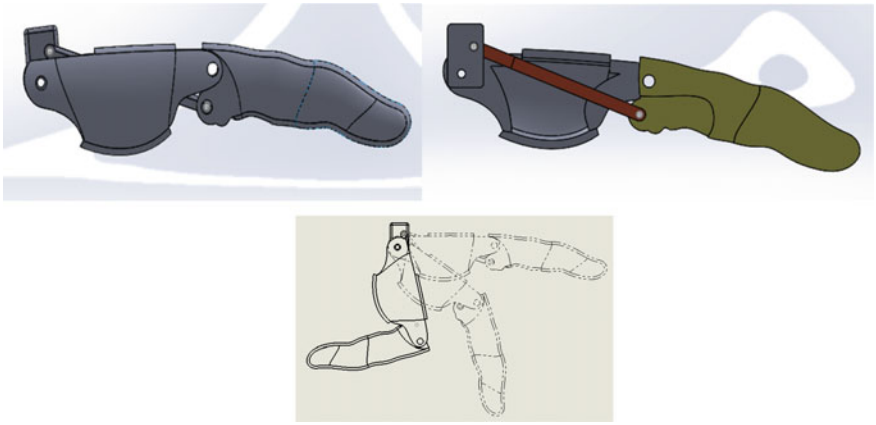


Fig. 1.21 Illustration of the kinematic model of the hand

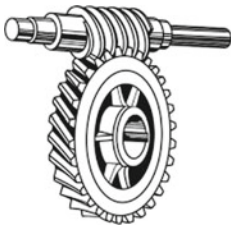


Fig. 1.22 Worm gear used as finger actuator

Due to size constraining, the actuator choice must be optimized with respect to the grasping force specifications.

The design should consider the availability of different product. Chinese product presents a good compromise between the torque of motors and their size, but their production can be corrupted at any time without a replacement. Other available products have a small size but does not meet torque requirements.

Faulhabert and Maxon, two companies, provide an interesting catalog of motors that meet the requirements for prosthetic hand, but expensive. Many prosthetic hands use their motors to actuate prosthetic hands, and they present a good quality and long-life cycle.

1.5 Results and Discussion

In the first stage, the classification is performed to the dataset after denoising the signal without any feature extraction. We used only the different envelope detected and compared which envelope will be used.

The table below shows the results of using SVM, KNN, ANN, and LDA to classify different signal envelope corresponding to hand motion of the five individuals.

The comparison is used to choose the best classifier and the best envelope detection method.

SVM uses one-versus-one strategy to classify data. The results of this classifier give a high accuracy using linear envelope detection method. To prevent an overfitting situation, we use a k-fold cross-validation to insure the stability and reliability of the given results.

The *k*-fold cross-validation consists of partitioning the original data to randomly *k* equal subsample. A single subsample is used to validate the model, and the rest *k*-1 subsample is used for validation.

KNN also uses the cross-validation technique with fivefold to ensure that no overfitting occurs. The results given below in figures and table illustrate the decrease in accuracy of the classifier with the increasing of the number of nearest neighbours. The results demonstrate that the accuracy decreases when the number of nearest neighbours increase (Fig. 1.23).

For LDA, we use pseudolinear version to classify the data. This classifier is characterized by its fast prediction, easy to interpret, and small memory usage.

Artificial neural networks give good results but present the inconvenient of slow prediction speed.

By comparing the results of the four-envelope detection method, we can clearly conclude that results of different classifier with peak envelope and linear envelope are the best, then, comes the RMS envelope, and finally the Hilbert envelope which is not suitable for this set of EMG signal (Table 1.2).

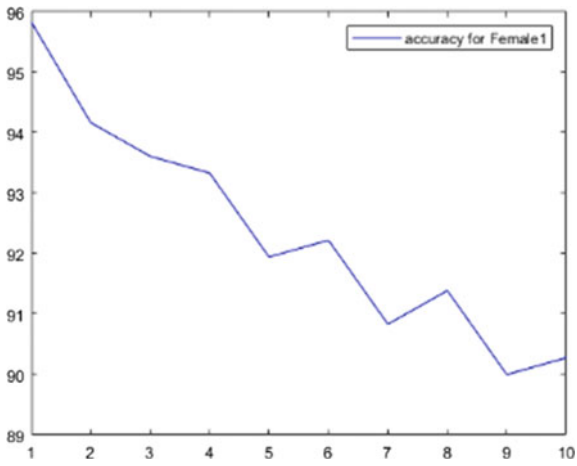


Fig. 1.23 Accuracy decreases when the number of nearest neighbours increase

Table 1.2 Classification accuracy for different envelope detection methods and different classifiers

Envelope classifier	Hilbert envelope	Peak envelope (%)	RMS envelope (%)	Linear envelope (%)	(%)
SVM (one-vs-one)	Male1	66.11	93.6	98.05	97.5
	Male2	44.16	78.8	96.9	95.0
	Female1	42.7	91.9	45.5	97.5
	Female2	69.72	93.3	70.27	98.3
	Female3	63.05	96.1	59.16	97.7
KNN (k = 1)	Male1	56.1	95.8	93.05	94.72
	Male2	38.5	91.1	80.0	90.27
	Female1	32.5	96.1	93.6	95.83
	Female2	48.8	94.17	67.2	95.27
	Female3	43.3	97.2	63.61	96.94
LDA	Male1	55.83	85.5	82.2	98.05
	Male2	32.7	89.4	75.27	97.2
	Female1	35.5	90.27	74.16	96.1
	Female2	57.77	81.6	81.1	94.4
	Female3	42.5	86.1	74.7	95.27
ANN	Male1	49.17	80.83	91.67	94.17
	Male2	40.83	75	58.33	68.33
	Female1	42.5	89.17	90.83	94.17
	Female2	58.33	81.67	90	90.83
	Female3	48.3	92.5	85	85.83

1.6 Conclusion

EMG signal are widely used in many applications where combining biological system with robotic system is preferred. In the field of prosthesis design, biological signals are the best way to control mechatronic systems. The evolution of hardware and software technology give the possibility obtain accurate results in the classification of those kind of signal. This work found that different stages of classification are essentially affected by the quality of the acquired signal. This fact can easily affect the identification results. The different types of noise presented in this chapter are the most known problems encountered when dealing with EMG signal processing. We also proved that envelope detection techniques are useful in the classification process and that peak envelope and linear envelope are the best method for extracting smooth and reliable envelopes. Those methods have an important impact in decreasing the noise deterioration of the general shape of the signal and thus keeping the information contained with those signals intact. Using feature extraction can also improve the classification results by neglecting the effect of noise by providing more stable

information and effectively discriminating classes. As discussed above, linear envelope and peak envelope give the best classification with our four classifiers SVM, KNN, ANN, and LDA. Future work will concentrate on building a real-time system for pre-processing and classifying EMG signal to use them effectively in real-world prosthetic hand. The stage of filtering of the signal will be the part of interest to increase the stability and the effectiveness of the system.

References

1. Sapsanis C, Georgoulas G, Tzes A (2013) "Improving EMG based classification of basic hand movements using EMD" in 35th annual international conference of the IEEE engineering in medicine and biology society '13. Osaka, Japan
2. Hale CA (2013) Physiotherapy for people with major amputation. Tidy's Physiotherapy. pp 457–474
3. Watve S, Dodd G, MacDonald R, Stoppard ER (2011) Upper limb prosthetic rehabilitation. Orthopaedics and Trauma 25(2):135–142
4. Thongpanja S, Phinyomark A, Quaine F, Laurillau Y, Limsakul C, Phukpattaranont P (2016) Probability density functions of stationary surface EMG signals in noisy environments. IEEE Trans Instrum Meas 65(7):1547–1557
5. De Luca CJ, Donald Gilmore L, Kuznetsov M, Roy SH (2010) Filtering the surface EMG signal: movement artifact and baseline noise contamination. J Biomech 43(8):1573–1579
6. Maddirala AK, Shaik RA (2016) Motion artifact removal from single channel electroencephalogram signals using singular spectrum analysis. Biomed Signal Process Control 30:79–85
7. De Luca CJ, Kuznetsov M, Gilmore LD, Roy SH (2012) Inter-electrode spacing of surface EMG sensors: reduction of crosstalk contamination during voluntary contractions. J Biomech 45(3):555–561
8. Gentilcore-Saulnier E., Auchincloss C., McLean L. (2016) Electromyography. In: Padoa A., Rosenbaum T. (eds) The Overactive Pelvic Floor. Springer, Cham
9. Kale S, Dudul S (2009) Intelligent noise removal from EMG signal using focused time-lagged recurrent neural network. Appl Comput Intell Soft Comput https://doi.org/10.1155/2009/129761
10. Mankar RV, Ghatol A (2008) Use of RBF neural network in EMG signal noise removal. p 7
11. Turnip A, Junaidi E (2014) Removal artifacts from EEG signal using independent component analysis and principal component analysis. 2014 2nd international conference on technology, informatics, management, engineering and environment, Bandung. pp 296–302
12. Zschorlich VR (1989) Digital filtering of EMG-signals. Electromyogr Clin Neurophysiol. 29(2):81–86 PMID: 2707144
13. Mello RGT, Oliveira LF, Nadal J (2007) Digital butterworth filter for subtracting noise from low magnitude surface electromyogram. Comput Methods Prog Biomed 87(1):28–35
14. Sapsanis C, Georgoulas G, Tzes A, (2013) EMG based classification of basic hand movements based on time-frequency features. pp 716–722. https://doi.org/10.1109/MED.2013.6608802
15. Comparison between envelope detection methods for bearing defects diagnose
16. Geethanjali P, Ray KK (2015) A low-cost real-time research platform for EMG pattern recognition-based prosthetic hand. IEEE/ASME Trans Mechatron 20(4):1948–1955
17. Rechy E, Hu H, (2018) Stages for developing control systems using EMG and EEG signals: a survey
18. Attenberger A, Buchenrieder K (2015) MATLAB/simulink-supported EMG classification on the raspberry Pi. In: Moreno-Díaz R, Pichler F, Quesada-Arencibia A (eds) Computer aided systems theory—EUROCAST 2015. EUROCAST 2015. Lecture Notes in Computer Science vol 9520. Springer, Cham

19. Riillo F, Quitadamo LR, Cavrini F, Gruppioni E, Pinto CA, Pastò NC, Sbernini L, Albero L, Saggio G (2014) Optimization of EMG-based hand gesture recognition: supervised versus unsupervised data pre-processing on healthy subjects and transradial amputees. Biomed Signal Process Control 14:117–125
20. Lotte F (2012) A new feature and associated optimal spatial filter for EEG signal classification: waveform length. In: Proceedings of the 21st international conference on pattern recognition (ICPR2012). Tsukuba, pp 1302–1305
21. Pancholi S, Joshi AM (15 April, 2018) Portable EMG data acquisition module for upper limb prosthesis application.IEEE Sens J 18(8):3436–3443
22. Krishnan S, Athavale Y (2018) Trends in biomedical signal feature extraction. Biomed Signal Process Control 43:41–63
23. Ju Z, Ouyang G, Wilamowska-Korsak M, Liu H (2013) Surface EMG based hand manipulation identification via nonlinear feature extraction and classification. IEEE Sens J 13(9):3302–3311
24. Gupta V, Mittal M (2018) KNN and PCA classifier with autoregressive modelling during different ECG signal interpretation. Procedia Comput Sci 125:18–24
25. Thongpanja S, Phinyomark A, Limsakul C, Phukpattaranont P (2015) Application of mean and median frequency methods for identification of human joint angles using EMG signal. In: Kim K (eds) Information science and applications. Lecture notes in electrical engineering vol 339. Springer, Berlin, Heidelberg
26. Portugal I, Alencar P, Cowan D (2018) The use of machine learning algorithms in recommender systems: a systematic review. Expert Syst Appl 97:205–227
27. Halilaj E, Rajagopal A, Fiterau M, Hicks JL, Hastie TJ, Delp SL (2018) Machine learning in human movement biomechanics: best practices, common pitfalls, and new opportunities. J Biomech
28. Tsoumakas G, Partalas I, Vlahavas I (2009) An ensemble pruning primer. In: Okun O, Valentini G (eds) Applications of supervised and unsupervised ensemble methods. Studies in computational intelligence vol 245. Springer, Berlin, Heidelberg
29. Nandy A, Biswas M (2018) Reinforcement learning basics. In: Reinforcement learning. Apress, Berkeley, CA
30. Yousefi J, Hamilton-Wright A (2014) Characterizing EMG data using machine-learning tools. Comput Biol Med 51:1–13
31. Purushothaman G, Ray KK (2014) EMG based man–machine interaction—a pattern recognition research platform. Robot Autonom Syst 62(6):864–870
32. Panchal-Kildare S, Malone K (2013) Skeletal anatomy of the hand. Hand Clin 29(4):459–471

Chapter 2
A Mobile Computing Solution for Enhanced Living Environments and Healthcare Based on Internet of Things

Gonçalo Marques

Abstract Since most people spend a considerable part of their time indoors, the indoor environment has a determining influence on human health. In several instances, the air quality parameters are extremely distinctive from those defined as healthy values. Using real-time monitoring, occupants or the build manager can administer interventions in order to improve indoor air quality (IAQ). The constant scientific improvements in numerous areas such as Ambient-Assisted Living and the Internet of Things (IoT) make it possible to build smart things with significant features for sensing and connecting. Therefore, the authors introduce an IoT architecture for real-time monitoring of IAQ. This system named ***iAQ Wi-Fi+*** uses an open-source Arduino UNO as processing unit, an ESP8266 for Wi-Fi 2.4 GHz as a communication unit, and incorporates a temperature and humidity sensor, a CO_2 sensor, a dust sensor, and a digital light sensor operating as a sensing unit. This solution is also composed of a smartphone application for data consulting. The monitored data can be discussed by clinicians to support medical diagnostics for enhanced healthcare. Compared to other solutions, the ***iAQ Wi-Fi+*** is based on open-source technologies and brings a Wi-Fi system, with several advantages such as its modularity, scalability, low-cost, and easy installation. The results obtained are very encouraging, representing a meaningful contribution to IAQ monitoring systems based on IoT.

2.1 Introduction

Indoor air quality (IAQ) assumes an important role as far as personal exposure to pollutants is concerned because several people such as retired, students in classrooms, and disabled persons could stay most of their time in indoor environments. Older people and new-borns who are most likely affected by poor IAQ may spend all their time in indoor environments [1].

G. Marques (✉)
Polytechnic of Coimbra, ESTGOH, Rua General Santos Costa, Oliveira do Hospital 3400-124, Portugal
e-mail: goncalosantosmarques@gmail.com

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_2

The assessment that IAQ indicators must thereby provide is how good and clean is it (a) by meeting thermal and respiratory requirements, (b) preventing unhealthy accumulation of pollutants, and (c) allowing for a sense of well-being [2].

Environmental Protection Agency (EPA) is responsible for indoor and outdoor air quality regulation in the USA. EPA considers that indoor levels of pollutants may be up to 100 times higher than outdoor pollutant levels and ranked poor air quality as one of the top 5 environmental risks to the public health [3].

The problem of poor IAQ becomes of utmost importance impacting severely the poorest people in the world who are most vulnerable presenting itself as a dangerous threat for world health such as tobacco use or the problem of sexually transmitted diseases [4].

Research should be done on IAQ to improve legislation and inspection and create real-time supervision systems for public health promotion. Those systems should be incorporated not only in public places such as institutions and hospitals but also in private places to increase the building's construction guidelines. Avoidance of smoking indoors and the use of natural ventilation are important behaviours which provide numerous positive impacts in IAQ and should be taught to children through educational programs [5]. Despite the importance of IAQ for public health, there is a lack of interest in the research community to study new methods for IAQ promotion [6].

The Ambient-Assisted Living (AAL) concept aims to create sophisticated tools and services to increase the quality of life, health, and well-being of older people by study new methods for enhanced living environments [7, 8].

However, some technical challenges still exist in the creation of AAL frameworks for enhanced occupational health [9]. Besides this, social and moral issues still exist, for example, the acceptance of these technologies by older persons and privacy and security problems which are remarkably imperative in each AAL project. It is essential to guarantee that innovation does not replace human care but rather introduces new methods to increase and support medical services and frameworks. At 2050, 20% of the total population will be aged 60 or above [10], which will increase the expansion and propensity for diseases, medical services costs, deficiency of parental figures, dependence, and a severer social effect. 87% of individuals wish to remain in their homes and bolster the enormous cost of nursing [11]. All these arguments assert the need to create systems to improve the quality of life for enhanced indoor living environments.

Internet of Things (IoT) stands as a standard where things are linked to the Internet and incorporate data collection capabilities. The basic idea of the IoT is the pervasive presence of a variety of objects with interaction and cooperation capabilities among them to reach a common objective [12–14].

This chapter aims at presenting the ***iAQ Wi-Fi+*** , a solution for IAQ monitoring based on IoT architecture. This solution is composed of a hardware prototype for ambient data collection and web/smartphone compatibility for data access. This system uses an open-source Arduino UNO as processing unit, an ESP8266 for Wi-Fi 2.4 GHz as communication unit and incorporates a temperature and humidity sensor, a CO_2 sensor, a dust sensor, and a digital light sensor as sensing unit. The previous

chapter presents a similar solution for wellness monitoring of the elderly by the observation of routing activities as an aid to caregivers. However, the ***iAQ Wi-Fi+*** uses an entirely wireless approach as well as providing easy installation.

The chapter is structured as follows: besides the introduction (Sect. 2.1), Sect. 2.2 introduces IoT; Sect. 2.3 is concerned to presents smart homes; Sect. 2.4 focuses on IAQ; Sect. 2.5 presents the materials and methods used in ***iAQ Wi-Fi+*** development; Sect. 2.6 shows the results, and the conclusions are in Sect. 2.7.

2.2 Internet of Things

The IoT concept is the result of different views: things-oriented vision, Internet-oriented vision and semantic-oriented vision [15]. The IoT consists of a global network of interconnected objects that have a unique address based on standard communication protocols. The things-oriented vision focuses on intelligent autonomous devices that use technologies such as near-field communication (NFC) and radio frequency identification (RFID) objects, applied to our daily lives. The Internet-oriented vision focuses on the idea of keeping the devices connected to the network, having a single address and using standard protocols.

Semantic vision focuses on storage, searching, and data classification generated by IoT. This vision centres on the development of software architecture solutions to manage the data produced by IoT devices.

The IoT is divided into five different layers such as Objects or Perception Layer, Object Abstract or Network Layer, Service Management or Middleware Layer, Application Layer, and Business Layer (Fig. 2.1).

On the one hand, the Perception Layer refers to physical sensors and actuators that IoT systems incorporate [16]. On the other hand, the Network Layer transfers data produced by the Perception Layer to the Middleware Layer through secure channels using technologies such as RFID, ZigBee, WPAN, WSN, DSL, UMTS,

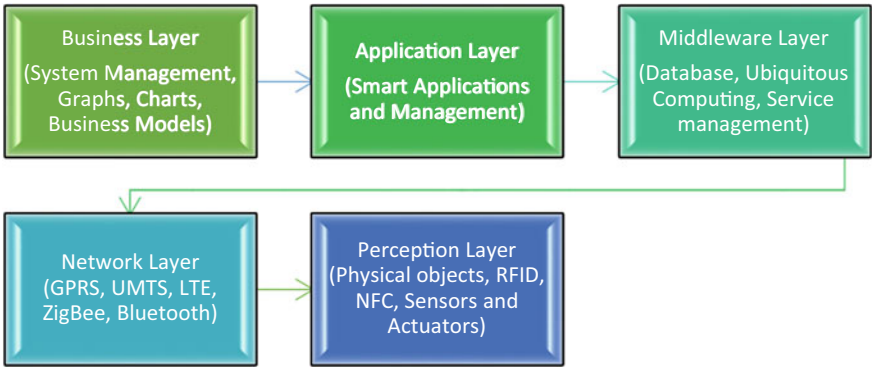


Fig. 2.1 IoT architecture layers (adapted from [16])

GPRS, Wi-Fi, WiMax, LAN, WAN, 3G, and LTE. Furthermore, the Middleware Layer pairs a service with its requester based on addresses and names to maintain independence from the hardware. On the contrary, the Application Layer provides the services requested by consumers passing the system output information to the client that demands that information.

Finally, the Business Layer manages the overall IoT system activities and services to build a business model, graphs and flowcharts based on the received data from the Application Layer.

Despite the numerous technologic enhancements, some issues in the construction of IoT systems continue to exist corresponding to privacy, confidentiality, and security of such systems. To address the privacy issues, the ambient sensing system (AmbLEDs) project presents the use of LEDs instead of other types of more invasive sensors such as cameras and microphones to interact with people in AAL revealing the importance of using intuitive interfaces in this applications [17].

Humans will often be an integral part of the IoT system. Therefore, the IoT will affect every aspect of human lives. Due to the large scale of devices carrying privacy and security issues, the cooperation between the research communities is deemed essential [18]. The SPHERE Project [19] aims to build a generic platform that fuses complementary sensor data to generate rich datasets that support the detection and management of various health conditions. This project uses three sensing technologies: environment, video and wearable sensing. SPHERE is a specially impressive project that makes a beneficial bridge between IoT and AAL. An AAL cloud-based IoT platform was proposed by [20] to manage the integration and behaviour-aware orchestration of devices as services stored and accessed via the cloud.

2.3 Smart Homes

There are decades of research about smart homes. The MIT Media Lab has developed the first project [21]. Today, three main categories of smart homes exist. The first category detects and identifies the actions of its residents to determine their health. The second category aims at the storing and retrieving of multimedia captured within the smart home, in different levels from photos to experiences. The third category is surveillance, where the data captured in the environment is processed to obtain information to raise alarms and protect the home and the residents. There is also a type of smart homes that have the objective to reduce energy consumption by monitoring and controlling electric devices [22].

Recent advances in information technology allowed lower prices of smart homes but provide them intelligence environments to make complex decisions remains a challenge. In the future, the number of smart homes will increase with the use of sensors that will store the data acquired in monitoring databases in real-time.

Three broad views are introduced by [23]: a functional view, an instrumental view, and a socio-technical view. The functional view sees smart homes as a way of better managing the demands of daily living through technology. The instrumental

view shows smart homes' potential for managing and reducing energy consumption. According to a socio-technical view, smart homes provide continuous digitalization to daily life activities.

In Europe, some smart home projects include iDorm [24], Gloucester Smart Home [25], CareLab [26], and others.

Several challenges are related to IoT and AAL such as security, privacy, and juridical. In general, IoT devices are wireless and exposed to a public range. Therefore, the ownership of data collected needs to be established. The IoT systems should adopt encryption methods and privacy policies.

An integrated platform for monitoring and control of a home that uses ZigBee wireless network and is distinguished by the use of open-source technologies that combines IoT and AAL has presented by [27].

As regular people spend a considerable part of their time inside buildings, the smart homes will have a significant role in occupational health. Therefore, smart homes should incorporate pervasive sensors for ambient quality evaluation to identify harmful situations in real-time and correlate the indoor environmental conditions with the occupant's health status.

2.4 Indoor Air Quality

An IAQ supervision system is an essential tool for air quality assessment, and furthermore is also an important decision tool for planning interventions to promote occupational health. An IAQ evaluation framework supports the identification and variation of IAQ parameters. Nearby and disseminated evaluation of IAQ is critical for enhanced occupational health, such as gas spills identification or security applications, but also to control heating, ventilation and air conditioning (HVAC) systems [6]. An IAQ real-time monitoring system gives a predictable stream of IAQ information for a stable administration and building managing. The remote IAQ monitoring provides continuous data collection for enhanced AAL standards, as proposed in [28], by presenting an effective air quality system for different gas sensors monitoring such as methane, propane, and carbon dioxide and monoxide.

In general, the existing IAQ frameworks are expensive and only implement random sampling. Thus, only a few new systems have been created for checking environmental parameters with the objective to enhance occupational health [29]. The accessibility of cost-effective sensors and communication technologies provide the criteria to develop intelligent systems for data collection and physical interaction, e.g., air quality monitoring and control systems [30]. Some solutions aimed at IAQ supervision are beginning to take centre stage [31–36]. A dedicated, miniaturized, low-cost electronic system based on metal oxide sensors and signal processing techniques that supports carbon monoxide, nitrogen dioxide in mixtures with relative humidity and volatile organic compounds supervision through an optimized gas sensor array and effective pattern recognition methods is presented by [37]. Another wireless solution for environmental parameters supervision such as temperature,

humidity, gaseous pollutants and aerosol is proposed by [38]. A monitoring system that uses a low-cost wireless sensor network, to collect IAQ information developed using Arduino, XBee modules and microsensors, for storage and availability of monitoring data in real-time is presented by [39].

2.5 Materials and Methods

In general, the quality of indoor environments covers the visual and thermal comfort but also the IAQ. In this context, environmental parameters such as temperature, humidity, airspeed, lighting level and pollutants concentrations are crucial for a proper evaluation of indoor environments.

The ***iAQ Wi-Fi+*** is an IAQ monitoring solution that provides temperature, humidity, PM_{10}, CO_2 and luminosity supervision in real-time. It is a completely wireless solution, where the wireless communication is implemented using the ESP8266 module which implements the IEEE 802.11 b/g/n networking protocol, a family of specifications developed by the IEEE for WLANs. The IEEE 802.11 standard supports radio transmission within the 2.4 GHz band [40].

This solution is based on open-source technologies, and it uses an Arduino UNO [41] as a microcontroller and an ESP8266 module for data communication. The data collected by the system is stored in a ThingSpeak platform. ThingSpeak is an open-source IoT application that offers APIs to store and retrieve data from sensors and devices using HTTP over the Internet [42].

The end user can access the data from the web page provided by ThingSpeak platform or can use the smartphone app developed in Swift, an open-source programming language with Xcode integrated development environment (IDE) created for the iOS operating system. By providing a history of changes, the system helps the build manager to make a precise and detailed analysis of the IAQ. Therefore, this data could support him to decide on possible interventions to improve IAQ.

Figure 2.2 represents the system architecture.

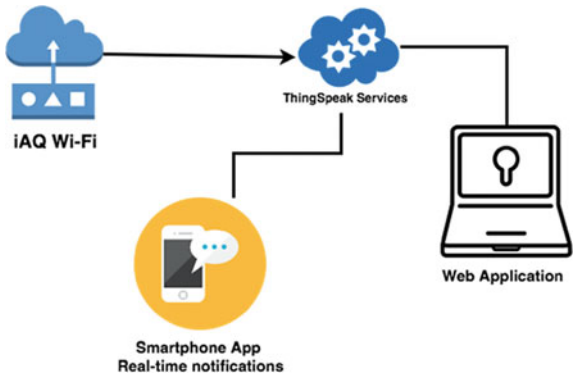


Fig. 2.2 ***iAQ Wi-Fi+*** system architecture

The ***iAQ Wi-Fi+*** can be divided into three major parts: a processing unit, a sensing unit, and a communication unit. This system is built using the embedded Arduino UNO microcontroller as the processing unit. Arduino is an open-source platform that incorporates an Atmel AVR microcontroller [41]. The sensing unit incorporates temperature, humidity, dust concentration, light, and CO_2 sensors. The system incorporates an ESP8266 as the communication unit (Fig. 2.3).

The ***iAQ Wi-Fi+*** prototype is shown in Fig. 2.4, and a brief description of the components used is described in Table 2.1.

The firmware of the ***iAQ Wi-Fi+*** is implemented using the Arduino platform language on the Arduino IDE. It belongs to the C-family programming languages. The Arduino UNO is responsible for data collection and is connected to the communication unit (ESP8266) by serial communication.

The Arduino UNO sends a string with the data from the sensors to the ESP8266, and this is responsible to upload this data to the ThingSpeak platform (Fig. 2.5).

Table 2.2 describes the cost of the components incorporated in the ***iAQ Wi-Fi+***.

Compared to the existent systems on the market which have an excessively higher cost and do not provide real-time monitoring data, the ***iAQ Wi-Fi+*** is a suitable solution for enhanced living environments and occupational health.

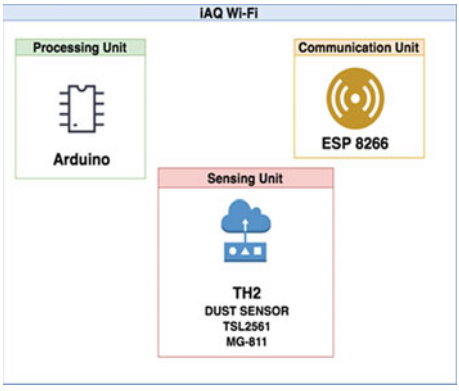


Fig. 2.3 Schematic diagram of ***iAQ Wi-Fi+***

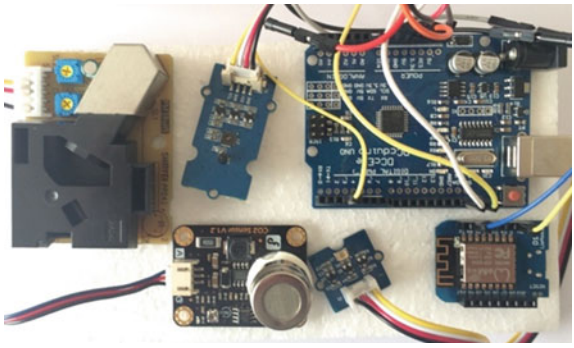


Fig. 2.4 ***iAQ Wi-Fi+*** prototype

Table 2.1 Type, interface, operating voltage and other details of the used components in the ***iAQ Wi-Fi+***

Component	Type	Interface	Operating voltage (V)	Details
Temperature and relative humidity sensor	TH2	I2C	3.3–5	Accuracy: ±4.5% RH and ±0.5 °C Range: 0–100%RH and 0–70 °C
Dust sensor	Shinyei Model PPD42NS	PWM	5	Sensitive to 1 μm and major particles [43]
Light sensor	TSL2561	I2C	5	Range: 0.1–40,000 lx Working temperature: (−40–85 °C) [44]
CO_2 sensor	MG-811	Analog	5	Range: 340–10,000 ppm [45]. Working temperature: 20–50 °C [46]
Wi-Fi chip and MCU (microcontroller unit)	ESP8266	I2C	3.3	32-bit MCU and supports 802.11 b/g/n protocols Working temperature: −40–125 °C [47]

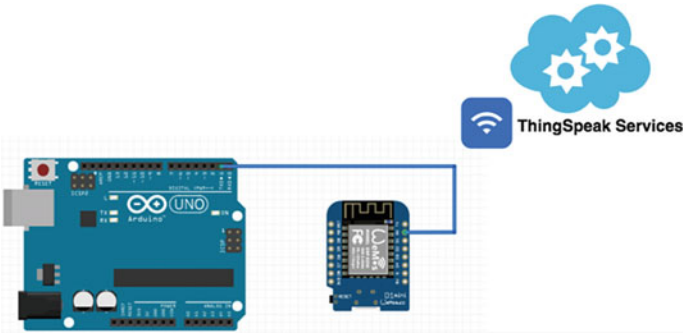


Fig. 2.5 Arduino UNO and ESP8266

Taking into consideration that ***iAQ Wi-Fi+*** is intended to be used in indoor environments where electricity is available, there was no great concern with the choice of ultra-low-power sensors and the research focus on the real-time data collection and notification features. The selection of the sensors was based on the cost of the system.

The ESP8266 has an important feature that provides to the end user an easy configuration of the Wi-Fi network to which it will be connected. The ESP8266 is by default a Wi-Fi client, but if it is unable to connect to the Wi-Fi network or if there

Table 2.2 Component costs of the ***iAQ Wi-Fi+*** system

	Cost
Arduino UNO	8.59 USD
ESP8266	3.39 USD
TH2	9.46 USD
Shinyei dust sensor	13.31 USD
TSL2561	11.20 USD
MG-811	39.10 USD
PCB	3.65 USD
Cables and box	9.59 USD
Total	98.29 USD

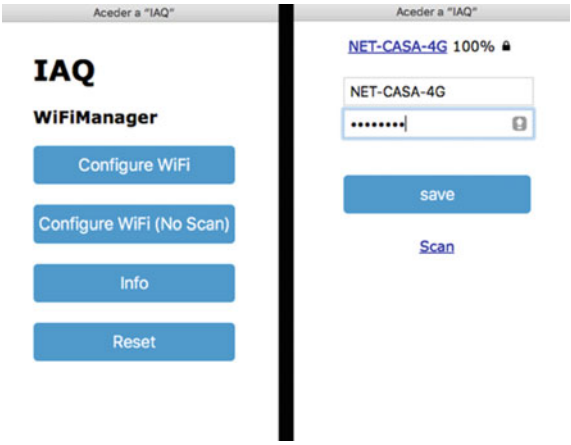


Fig. 2.6 Wi-Fi network configuration

are no wireless networks available, the ESP8266 will turn to hotspot mode and will create a Wi-Fi network with an SSID "IAQ". At this point, the end user can connect to the created hotspot which permits the configuration of the Wi-Fi network to which the ***iAQ Wi-Fi+*** is going to connect through the introduction of the network SSID and password (Fig. 2.6).

2.6 Results and Discussion

The ***iAQ Wi-Fi+*** allows viewing the data according to graphical and numerical values by using a web browser or a smartphone app.

For testing purposes, two ***iAQ Wi-Fi+*** modules were been used. All modules are powered using 230–5 V AC-DC 2A power supply. IAQ data were collected for two months and show that under certain conditions, air condition is significantly affected.

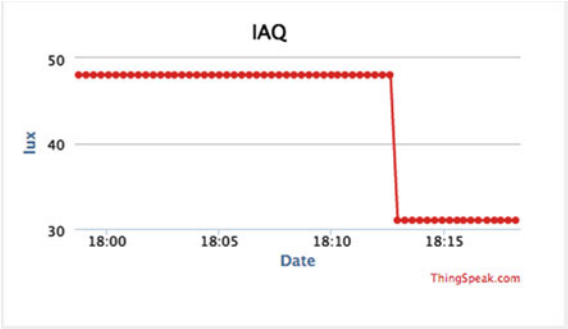
The end user can access the data using the ThingSpeak platform but also from the web portal built in PHP. After login, the end user can easily access the IAQ data in real-time. The web application allows the user to keep the parameters history. The system helps the user to provide a precise and detail analysis of the air quality behaviour. The map view feature allows the user to check in real-time the latest values collected by ***iAQ Wi-Fi+*** referencing their location (Fig. 2.7).

A sample of the data collected by ***iAQ Wi-Fi+*** is shown in Figs. 2.8, 2.9, 2.10 and 2.11. Figure 2.8 represents luminosity data measured in lux, Fig. 2.9 represents temperature data measured in Celsius, Fig. 2.10 represents dust sensor data measured in $\mu g/m^3$ and Fig. 2.11 represents the humidity data measured in %. The graphs display the results obtained in a real environment with induced simulations.

The ***iAQ Wi-Fi+*** is also equipped with a powerful alerts manager that notifies the user when the IAQ is poor. The maximum and minimum health quality values are

Fig. 2.7 Map view functionality

Fig. 2.8 Luminosity (lux)



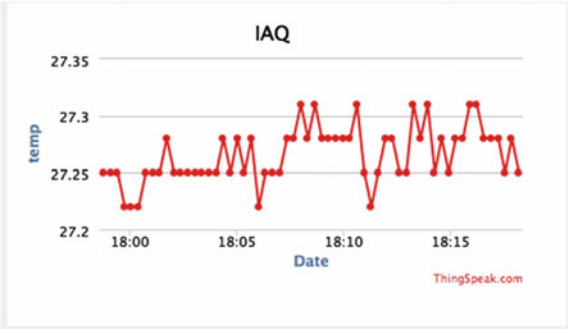


Fig. 2.9 Temperature (ºC)

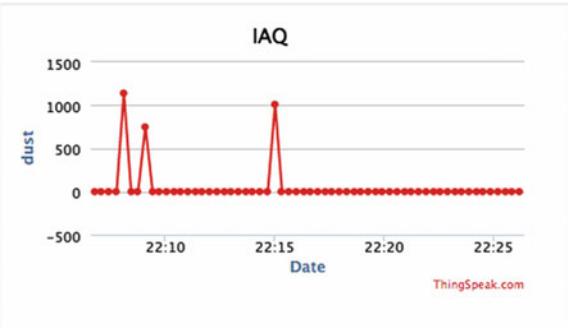


Fig. 2.10 Dust sensor PM_{10} ($\mu g/m^3$)

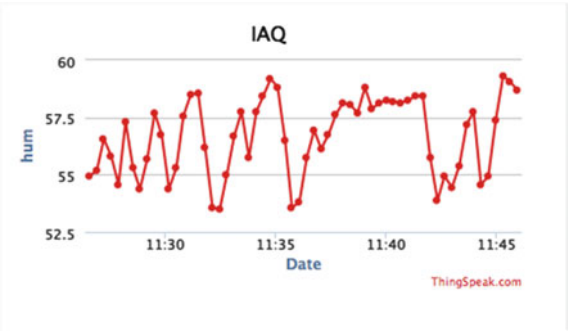


Fig. 2.11 Humidity (%)

predefined by the system based on well-studied values, but the user can also change these values to specific proposes (Fig. 2.12).

When a value exceeds the defined threshold, the user will be notified in two ways by e-mail or by smartphone notification in real-time (Fig. 2.13). This functionality enables the user to act in real-time ensuring excellent ventilation for enhanced living environments and occupational health.

The smartphone application allows a quick, simple, intuitive and real-time access to the monitored data in numerical and graphical form (Fig. 2.14). Mobile computing

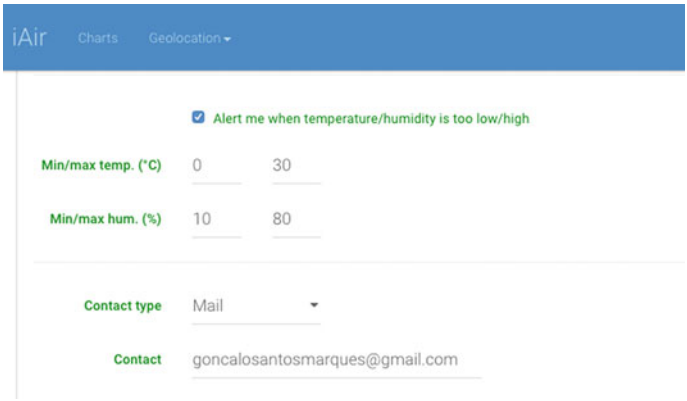


Fig. 2.12 Alert configuration manager

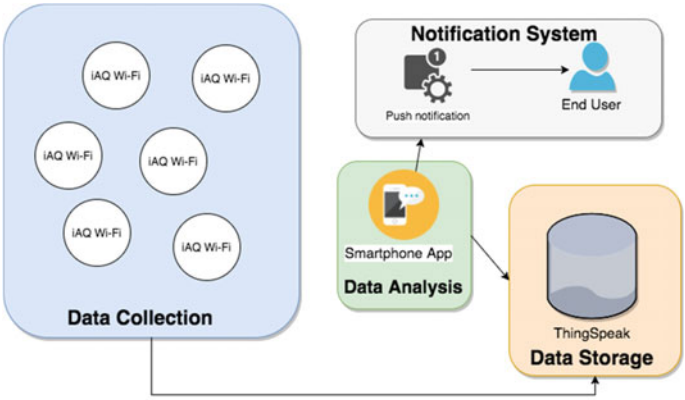


Fig. 2.13 ***iAQ Wi-Fi+*** smartphone notifications architecture

in the USA has an exponential growth as adult smartphone device ownership was at 33% in 2011, 56% at the end of 2013 and 64% in early 2015 [48]. In the Netherlands, 70% of the global population and 90% of the adolescents own a smartphone [49]. In Germany, 40% of the population uses a smartphone [50] and 51% of adults owned smartphones in the UK [51]. About 36–40% of smartphone owners use their smartphone 5 min before bed and in the next 5 min after waking up [50]. Smartphones not only have excellent processing and area capabilities but also people carry them in their daily lives. Therefore, a mobile application has been created to allow a quick, easy and intuitive access to the monitoring data. In this way, the user can carry the IAQ data of their home with him for everyday use.

On the one hand, the IAQ charts allow greater perception of the parameters behaviour than the numerical format. On the other hand, mobile computing technologies enable a precise analysis of the temporal evolution. Thus, the system is

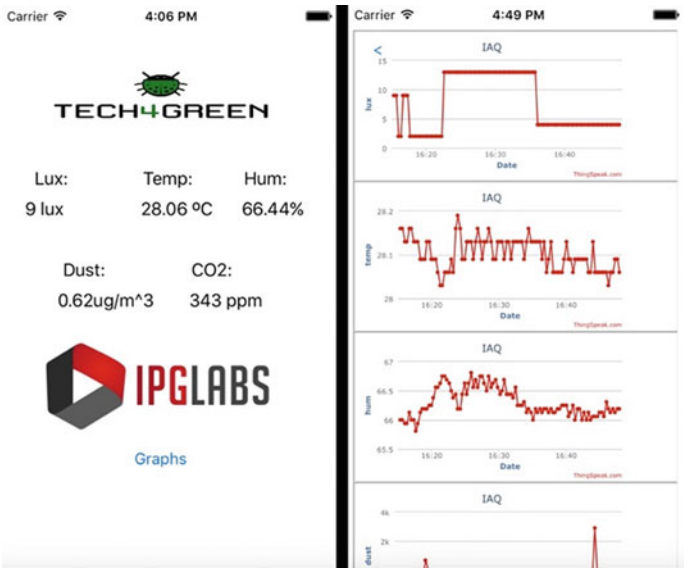


Fig. 2.14 ***iAQ Wi-Fi+*** mobile application

a powerful tool for air quality evolution analysis and decision making on possible interventions for enhanced living environments and occupational health.

Compared to other systems, the ***iAQ Wi-Fi+*** system has the following advantages: modularity, small size, low-cost construction and easy installation [52, 53].

In the future, the principal objective is to improve the work by adding a context-awareness approach to automate the of HVAC equipment's [54–57]. Improvements to the system hardware and software are planned to make it much more appropriate for specific purposes such as hospitals, schools, and offices. Formaldehyde (FA) is a high-volume chemical, which is used for disinfection purposes and as a preservative. FA is genotoxic, causing DNA adduct formation, and has a clastogenic effect; exposure–response relationships were nonlinear [58]. Therefore, the possibility to incorporate an FA sensor in ***iAQ Wi-Fi+*** should be noted.

2.7 Conclusion

By monitoring IAQ is possible to provide analysis of the ventilation conditions in real-time and plan interventions to increase the air quality if needed.

This chapter has presented a complete wireless solution for IAQ monitoring based on IoT architecture. This solution is composed of a hardware prototype for ambient data collection and smartphone compatibility for data access. The results obtained are promising, representing a significant contribution to IAQ monitoring systems based on IoT.

Compared to existing systems, it has great potential due to the use of low-cost and open-source technologies. Note that the system has advantages for both easy installation and configuration, due to the use of wireless technology for communications, but also because it was meant to be compatible with all domestic house devices and not only for smart houses or high-tech houses.

Despite all the advantages in the use of IoT architecture, still exist many open issues as scalability, quality of service problems, and security and privacy issues. The proposed system should find ways to respond to these problems. Data security, in particular, is of the utmost importance in the field of health care. Therefore, Part II of this book presents several significant approaches on security and privacy in IoT-based e-health applications. As future work, it is expected to introduce new sensors to this system for monitoring other IAQ parameters as well as the development of a platform that allows sharing in a secure way the collected data to health professionals. In addition to system validation, physical system and related software improvements have been planned to adjust the system to specific cases such as hospitals, schools, and industrial factories.

Systems like ***iAQ Wi-Fi+*** will be a part of the indoor spaces in the future. IAQ data can be used by health professionals to support clinical diagnostics to address occupational health problems.

References

1. Diffey B (2018) Time and place as modifiers of personal UV exposure. Int J Environ Res Public Health 15(6):1112
2. Kumar SR, Davies S, Weitzman M, Sherman S (2015) A review of air quality, biological indicators and health effects of second-hand waterpipe smoke exposure. Tobacco Control 24(Suppl 1):i54–i59
3. Seguel JM, Merrill R, Seguel D, Campagna AC (2016) Indoor Air Quality. Am J Lifestyle Med 11(4):284–2895
4. Bruce N, Pope D, Rehfuess E, Balakrishnan K, Adair-Rohani H, Dora C (2015) WHO indoor air quality guidelines on household fuel combustion: strategy implications of new evidence on interventions and exposure–risk functions. Atmos Environ 106:451–457
5. Butz AM (2011) A randomized trial of air cleaners and a health coach to improve indoor air quality for inner-city children with Asthma and secondhand smoke exposure. Arch Pediatr Adolesc Med 165(8):741
6. De Vito S et al (2011) Cooperative 3D air quality assessment with wireless chemical sensing networks. Procedia Eng 25:84–87
7. Colantonio S, Coppini G, Giorgi D, Morales M-A, Pascali MA (2018) Computer vision for ambient assisted living. In: Computer vision for assistive healthcare. Elsevier, pp 147–182
8. Marques G (2019) Ambient assisted living and internet of things. In: Cardoso PJS, Monteiro J, Semião J, Rodrigues JMF (eds) Harnessing the internet of everything (IoE) for accelerated innovation opportunities. IGI Global, Hershey, PA, USA, pp 100–115
9. Koleva P, Tonchev K, Balabanov G, Manolova A, Poulkov V (2015) Challenges in designing and implementation of an effective ambient assisted living system. In: Telecommunication in modern satellite, cable and broadcasting services (TELSIKS), 2015 12th international conference on. pp 305–308
10. McNicoll G (2002) World population ageing 1950–2050. Popul Dev Rev 28(4):814–816

11. Centers for disease control and prevention, The state of aging and health in America 2007. N. A. on an aging society, 2007, Available: https://www.cdc.gov/aging/pdf/saha_2007.pdf
12. Giusto D (ed) (2010) The internet of things: 20th tyrrhenian workshop on digital communications. Springer, New York
13. Marques G, Pitarma R (2017) Monitoring and control of the indoor environment. In: 2017 12th Iberian conference on information systems and technologies (CISTI). pp 1–6.
14. Marques G, Pitarma R (2017) Monitoring energy consumption system to improve energy efficiency. In: Rocha Á, Correia AM, Adeli H, Reis LP, Costanzo S (eds) Recent advances in information systems and technologies, vol 570. Springer International Publishing, Cham, pp 3–11
15. Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. Comput Netw 54(15):2787–2805
16. Al-Fuqaha A, Guizani M, Mohammadi M, Aledhari M, Ayyash M (2015) Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun Surv Tutorials 17(4):2347–2376
17. Cunha M, Fuks H (2014) AmbLEDs para ambientes de moradia assistidos em cidades inteligentes," In: Proceedings of the 13th Brazilian symposium on human factors in computing systems, Foz do Iguaçu. Brazil, pp 409–412
18. Stankovic JA (2014) Research directions for the internet of things. Int Things J IEEE 1(1):3–9
19. Zhu Ni et al (2015) Bridging e-health and the internet of things: the SPHERE project. Intell Syst IEEE 30(4):39–46
20. Cubo J, Nieto A, Pimentel E (2014) A cloud-based internet of things platform for ambient assisted living. Sensors 14(8):14070–14105
21. Moukas A, Zacharia G, Guttman R, Maes P (2000) Agent-mediated electronic commerce: an MIT media laboratory perspective. Int J Electron Commer 4(3):5–21
22. De Silva LC, Morikawa C, Petra IM (2012) State of the art of smart homes. Eng Appl Artif Intell 25(7):1313–1321
23. Wilson C, Hargreaves T, Hauxwell-Baldwin R (2015) Smart homes and their users: a systematic analysis and key challenges. Pers Ubiquitous Comput 19(2):463–476
24. Pounds-Cornish A, Holmes A (2002) The iDorm—a practical deployment of grid technology. In: Cluster Computing and the Grid, 2002. 2nd IEEE/ACM International Symposium on. pp 470–470
25. Orpwood R, Gibbs C, Adlam T, Faulkner R, Meegahawatte D (2004) The gloucester smart house for people with dementia—user-interface aspects. In: Keates S, Clarkson J, Langdon P, Robinson P (eds) Designing a more inclusive world. Springer, London, pp 237–245
26. Henkemans OB, Caine KE, Rogers WA, Fisk AD (2007) Medical monitoring for independent living: user-centered design of smart home technologies for older adults. In: Processing med-e-tel conference eHealth, telemedicine and health information and communication technologies. pp 18–20
27. Suryadevara NK, Kelly S, Mukhopadhyay SC (2014) Ambient assisted living environment towards internet of things using multifarious sensors integrated with XBee Platform. In: Mukhopadhyay SC, (ed) Internet of things vol 9, Springer International Publishing. pp 217–231
28. Preethichandra DMG (2013) Design of a smart indoor air quality monitoring wireless sensor network for assisted living. pp 1306–1310
29. Yu T-C, Lin C-C (2015) An intelligent wireless sensing and control system to improve indoor air quality: monitoring, prediction, and preaction. Int J Distrib Sens Netw 11(8):140978
30. Al-Haija QA, Al-Qadeeb H, Al-Lwaimi A (2013) Case study: monitoring of AIR quality in king faisal university using a microcontroller and WSN. Procedia Comput Sci 21:517–521
31. Marques G, Roque Ferreira C, Pitarma R (2018) A system based on the internet of things for real-time particle monitoring in buildings. Int J Environ Res Publ Health 15(4):821
32. Marques G, Pitarma R (2018) IAQ evaluation using an IoT CO_2 monitoring system for enhanced living environments. In: Rocha Á, Adeli H, Reis LP, Costanzo S (eds) Trends and advances in information systems and technologies, vol 746. Springer International Publishing, Cham, pp 1169–1177

33. Marques G, Pitarma R (2019) Smartwatch-based application for enhanced healthy lifestyle in indoor environments. In: Omar S, Haji Suhaili WS, Phon-Amnuaisuk S (eds) Computational intelligence in information systems, vol 888. Springer International Publishing, Cham, pp 168–177
34. Pitarma R, Marques G, Caetano F (2016) Monitoring indoor air quality to improve occupational health. In: Rocha Á, Correia AM, Adeli H, Reis LP, Mendonça Teixeira M (eds) New advances in information systems and technologies, vol 445. Springer International Publishing, Cham, pp 13–21
35. Marques G, Pitarma R (2019) A cost-effective air quality supervision solution for enhanced living environments through the internet of things. Electronics 8(2):170
36. Marques G, Ferreira CR, Pitarma R (Mar. 2019) Indoor air quality assessment using a CO_2 monitoring system based on internet of things. J Med Syst 43(3)
37. Zampolli S et al (2004) An electronic nose based on solid state sensor arrays for low-cost indoor air quality monitoring applications. Sens Actuators B: Chem 101(1–2):39–46
38. Bhattacharya S, Sridevi S, Pitchiah R (2012) Indoor air quality monitoring using wireless sensor network. pp 422–427
39. Marques G, Pitarma R (2016) Health informatics for indoor air quality monitoring. In: Information systems and technologies (CISTI), 11th Iberian conference on, 2016. pp 1–6
40. Bhoyar R, Ghonge M, Gupta S (2013) Comparative study on IEEE standard of wireless LAN/Wi-Fi 802.11 a/b/g/n. Int J Advan Res Electron Commun Eng (IJARECE) 2(7)
41. D'Ausilio A (2012) Arduino: a low-cost multipurpose lab equipment. Behav Res Methods 44(2):305–313
42. Doukas C, Maglogiannis I (2012) Bringing IoT and cloud computing towards pervasive healthcare. pp 922–926
43. Austin E, Novosselov I, Seto E, Yost MG (2015) Laboratory evaluation of the Shinyei PPD42NS low-cost particulate matter sensor. PLoS ONE 10(9):e0137789
44. Minghui Y, Peng Y, Wangwang S (2010) Light intensity sensor node based on TSL2561 [J]. Microcontrollers Embed Syst 6:017
45. da Lima AL, da Silva VL (2015) Micro sensor para monitoramento da qualidade do ar. In: Workshop de Gestão, Tecnologia Industrial e Modelagem Computacional. vol 1
46. Banick JL, Zolkowski JJ, Lenz KE, Sanders J (2016) Monitoring carbon dioxide and methane levels above retired landfill and forest control site with a tethered aerostat to determine remediation effectiveness. In: Proceedings of the wisconsin space conference
47. Espressif Systems (2015) "ESP8266EX Datasheet," https://download.arduino.org/products/UNOWIFI/0A-ESP8266-Datasheet-EN-v4.3.pdf
48. Müller H, Gove JL, Webb JS, Cheang A (2015) Understanding and comparing smartphone and tablet use: insights from a large-scale diary study. In: Proceedings of the annual meeting of the australian special interest group for computer human interaction. pp 427–436
49. van Deursen AJAM, Bolle CL, Hegner SM, Kommers PAM (2015) Modeling habitual and addictive smartphone behavior. Comput Hum Behav 45:411–420
50. Montag C et al (Dec. 2015) Smartphone usage in the 21st century: who is active on WhatsApp?. BMC Res Notes 8(1)
51. Pearson C, Hussain Z (2015) Smartphone use, addiction, narcissism, and personality: a mixed methods investigation. Int J Cyber Behav, Psychol Learn 5(1):17–32
52. Marques G, Pitarma R (2016) An indoor monitoring system for ambient assisted living based on internet of things architecture. Int J Environ Res Public Health 13(11):1152
53. Pitarma R, Marques G, Ferreira BR (Feb. 2017) Monitoring indoor air quality for enhanced occupational health. Journal of Medical Systems 41(2)
54. Lima WS, Souto E, Rocha T, Pazzi RW, Pramudianto F (2015) User activity recognition for energy saving in smart home environment. In: 2015 IEEE symposium on computers and communication (ISCC), Larnaca. pp 751–757
55. Najem N, Ben Haddou D, Abid MR, Darhmaoui H, Krami N, Zytoune O (2017) Context-aware wireless sensors for IoT-centeric energy-efficient campuses. In: 2017 IEEE international conference on smart computing (SMARTCOMP). Hong Kong, China pp 1–6

56. Wu C-L, Tseng Y-S, Fu L-C (2013) Spatio-temporal feature enhanced semi-supervised adaptation for activity recognition in IoT-based context-aware smart homes. In: 2013 IEEE international conference on green computing and communications and IEEE internet of things and IEEE cyber, physical and social computing. Beijing, China. pp 460–467
57. Lohani D, Acharya D (2016) "SmartVent: a context aware IoT system to measure indoor air quality and ventilation rate. In: 2016 17th IEEE international conference on mobile data management (MDM), Porto. pp 64–69
58. Nielsen GD, Larsen ST, Wolkoff P (2017) Re-evaluation of the WHO (2010) formaldehyde indoor air quality guideline for cancer risk assessment. Arch Toxicol 91(1):35–61

Chapter 3
Rapid Medical Images Restoration Combining Parametric Wiener Filtering and Wave Atom Transform Based on Local Adaptive Shrinkage

Zouhair Mbarki and Hassene Seddik

Abstract Supervised image restoration is a process of reconstructing or recovering an image that has been degraded by using a priori knowledge of the degradation phenomenon. This includes deblurring images degraded by the limitations of sensors or source of captures, in addition to noise filtering and correction of geometric distortion due to sensors. Generally, restoring original images from their distorted form is a necessary step in many domains such as medical domain. In the literature, there are several classical medical images restoration techniques such as Wiener filtering. Thus, restoration techniques are oriented toward modeling the degradation and applying the inverse process in order to find an estimate of the original image. To recover the original medical image, Wiener filter needs a prior knowledge of the degradation phenomenon caused by the imaging system, the blurred image, and the statistical properties of the noise process. In this work, we propose a new fast algorithm for supervised medical images restoration that does not require a priori knowledge about the noise distribution. The degraded image is first deconvoluted in Fourier space by parametric Wiener filtering, then it is smoothed by the wave atom transform after applying a local adaptive shrinkage to its coefficients. Experiment results are very interesting and show the efficiency of the suggested method based on a comparison study.

3.1 Introduction

Medical images restoration remains a challenging task in the field of image processing. During the last decades, medical systems imaging has been developing and many techniques are available like computed tomography (CT), ultrasonography, magnetic resonance imaging (MRI), and positron emission tomography (PET).

Z. Mbarki (✉) · H. Seddik
Department of Electrical Engineering, CEREP, ENSIT, University of Tunis, 5, Av. Taha Hussein, 1008 Tunis, Tunisia
e-mail: mbarki.zouhair84@gmail.com

H. Seddik
e-mail: seddik-hassene@ieee.org

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_3

Despite the performance of these systems, often, medical images have a major disadvantage. Indeed, acquired images are degraded during the imaging process. Image degradations are, on the one hand, linked to the noise caused by the sensor, the quantification, and the transmission, on the other hand, are convolutive. These degradations are related to the movement of the sensor or the movement of the patient, and they are interpreted by the presence of the blur in the image. This image will be characterized by its spread function (PSF). Typically, the distorted image is caused by a degradation function denoted $H(x,y)$ which is generally a PSF followed by noise denoted $\eta(x,y)$. In order to make the deteriorate images exploitable and facilitate their analysis, a processing operation is necessary. This operation is called image restoration. According to the information about the degradation function (PSF) and the phenomenon of noise, the restoration is called supervised or blind restoration. Supervised image restoration, often called classical restoration, seeks an estimate of the true image assuming the blur is known. It attempts to reconstruct or recover an image that has been degraded by using a priori knowledge of the mathematical degradation model. Thus, restoration techniques focus on modeling the degradation and applying the inverse model in order to recover a denoised image. A fundamental method in the filtering theory used commonly for image restoration is the Wiener filter. The goal of the Wiener filter is to compute a statistical estimate of unknown signal using a related signal as an input and filtering it to generate the estimate as output. This method's objective is to restore image as close as possible to the original one. This is, indeed, a method of least squares estimation in which we consider that the images are realizations of a stationary and ergodic random process. The drawback of this method is the need for a priori knowledge of the degradation phenomenon, which is denoted as the degradation function of the imaging system, i.e., the point spread function (PSF), the blurred image, and the statistical properties of the noise process. To restore distorted medical image and improve Wiener filtering without prior knowledge of the degradation phenomenon, several authors propose to estimate the PSF using the specification of the imaging system or ground characters in the acquired images. Then, an appropriate method is selected to restore the images with the estimated PSF [1, 2]. In other works, authors proposed to restore images in the wavelet domain [3, 4]. Donoho and Johnstone in [5] proposed an algorithm for solving the inverse problem known as the wavelet-vaguelletes algorithm. It consists of applying an inverse filter and then the wavelet transform. Recently, a fast algorithm (VSFIR) for image restoration has been proposed. This approach is based on a variable splitting to obtain an equivalent constrained optimization formulation which is addressed with an augmented Lagrangian method [6]. In other work [7], authors proposed a supervised algorithm for medical images restoration using the conceptual framework of multiple-point geostatistics. In [8], a new frame for image estimation is provided. In fact, authors proposed the AFO TV-$\ell 1$ model with a novel IFODKM model. The proposed method is resolved by using SBI algorithm. In [9],

a TV blind method for single frame by employing split Bregman iteration called as TVBDSB is presented. The proposed method can excellently recover the degraded images not only with simple background but also with complex background. In [10], a non-local regularization algorithm with a single forward backward operator splitting is proposed to solve the sub-problems of the Bregman iterations. In related work [11], authors proposed a non-blind image restoration method (NBID) that combines the TV and NLTV models. Firstly, original image is decomposed into three regions, salient edges, details, and constant regions, and then the TV model and the NLTV model are used to smooth the image and preserve the salient edges and details, respectively. Other techniques propose to estimate the parameters of the PSF from the degraded image. This is achieved by training ELM using geometric moments as feature vectors [12]. In the present study, in order to benefit from the advantages of the Fourier transform and those of wave atom, we propose a scheme that consists of two steps: Deconvolution in the Fourier domain followed by a denoising in the wave atom domain. The threshold is calculated locally by estimating the noise variance in each region after the image segmentation operation is deployed using watershed algorithm. The simulations results and the comparative study with other techniques are conducted, and they have shown that the proposed schema generates a wide improvement in the quality of the restored images. This improvement is accounted subjectively in terms of visual quality and objectively with reference to the computation of some criteria.

3.2 Image Degradation Process

3.2.1 Notion of Blur

The digital filters used for images processing will have different effects according to their characteristics (data of the matrix). Some filters will enhance the edge effect, to give more apparent clarity to the image, others will have the opposite effect, to generate blur which is essentially a convolutive phenomenon and is characterized by its point spread function (PSF).

On the one hand, the blur is due to the sensor, and on the other hand to the shooting conditions.

- **Sensor defects**: Basically, the blur is due to the optics and the integration of the light on each pixel, the electronic noise, the quantification noise, and the replacement of the pixel size by the grain size on the analogical carriers.
- **External Defects**: Blur is caused by camera motion and atmospheric disturbances.

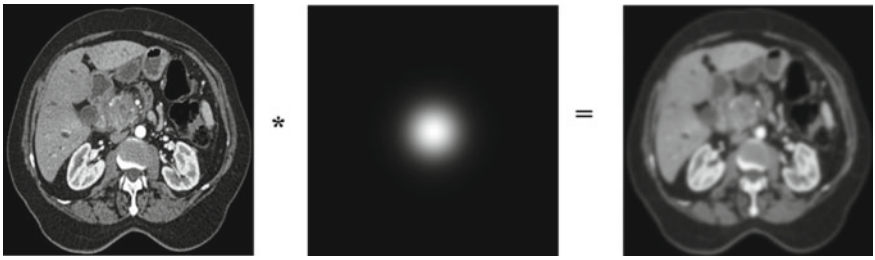


Fig. 3.1 Gaussian blur ($\sigma = 2$)

3.2.2 Various Types of Image Blur

3.2.2.1 Gaussian Blur

The Gaussian blur is a type of image-blurring filter that uses a Gaussian function, which also expresses the normal distribution in statistics, to calculate the transformation to be applied to each pixel in the image. The equation of a Gaussian functions in two dimensions is governed by Eq. (3.1).

$$G(x, y) = \frac{2}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.1}$$

where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and σ is the standard deviation of the Gaussian distribution.

The term $\frac{1}{\sigma\sqrt{2\pi}}$ is the normalization constant. It comes from the fact that the integral on the exponential function is not the unit:

$$\int_{-\infty}^{+\infty} e^{\frac{-x^2}{2\sigma^2}} \mathrm{d}x = \sqrt{2\pi}\sigma \tag{3.2}$$

With the normalization constant, the Gaussian kernel is a normalized kernel. Its integral over its entire domain is the unit for each value of σ. This means that increasing the value of σ substantially reduces the amplitude. Figure 3.1 shows an example of Gaussian blur.

3.2.2.2 Motion Blur

Motion blur refers to the image degradation caused by system movement or shake during the imaging process. Geometrical optic analysis shows that the point spread function (PSF) of motion blur is a uniformly distributed circular facula. An example of motion blur is shown in Fig. 3.2.

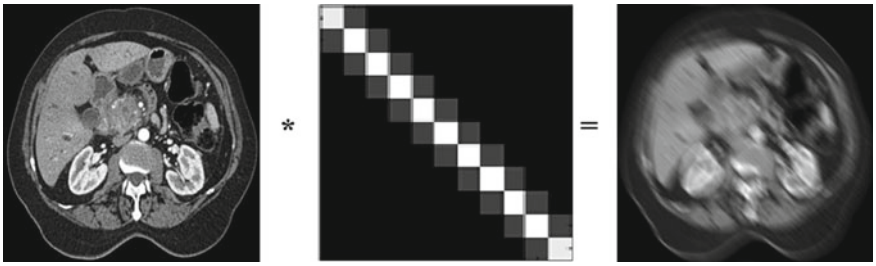


Fig. 3.2 Motion blur (Len = 15 et θ =135°)

3.2.2.3 Average Blur

This blur is produced when the camera is not in focus. Indeed, this filter assigns to the studied pixel the mean value of this pixel and the adjacent pixels. Figure 3.3 shows an example of average blur.

3.2.3 Mathematical Formulation of Image Degradation

In practice, each imaging system is characterized by its impulse response or point spread function (PSF) and assuming that the noise is additive and that it is uncorrelated with the signal. The discrete model for a linear degradation caused by blurring and additive noise is given by Eq. (3.3).

$$g(x, y) = H[f(x, y)] + \eta(x, y) \tag{3.3}$$

where $g(x, y)$ is the observed image, $H(x, y)$ is the degradation matrix, $f(x, y)$ is the original image, and $\eta(x, y)$ is a measurement noise.

The image restoration has for object the reduction see the elimination of the distortions introduced by the system having served to acquire the image. Its objective

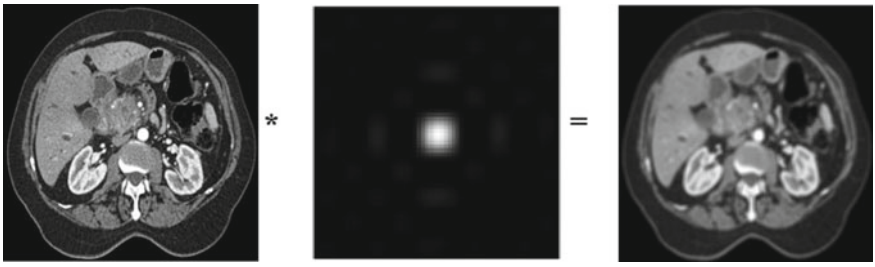


Fig. 3.3 Average blur

is to obtain an estimate $\hat{f}(x, y)$ of the original image. The estimated image must be as close as possible to the original one, and the error between those images should not exceed the Weber threshold and must be imperceptible psycho-visually. If the degradation function called $H(x, y)$ is a linear spatially-invariant process, it can be shown that the degraded image is given in the spatial domain by the following equation:

$$g(x, y) = h(x, y) * f(x, y) + \eta(x, y) \tag{3.4}$$

where $f(x, y)$ is the original image, $g(x, y)$ is the observed image, $h(x, y)$ is the blur kernel or PSF, $\eta(x, y)$ is the noise, and "$*$" indicates the convolution operator. Assuming that convolution in the spatial domain and multiplication in the frequency domain constitute a Fourier transform pair, we can write the preceding model in an equivalent frequency domain representation.

$$G(u, v) = H(u, v) \cdot F(u, v) + N(u, v) \tag{3.5}$$

where capital letters represents the Fourier transform of the corresponding terms in the spatial domain. The degradation function $H(u, v)$ is sometimes called the optical transfer function (OTF), a term derived from the Fourier analysis of optical systems. In the spatial domain, $h(x, y)$ is referred to as the point spread function (PSF) and operates on a point of light to obtain the characteristics of the degradation for any type of input. Because the degradation is due to a linear space-invariant function $H(x, y)$, which can be modeled as convolution, the degradation process is sometimes referred to as "convolving the image with PSF." Similarly, the restoration process is sometimes referred to as deconvolution. Figure 3.4 shows the image degradation process.

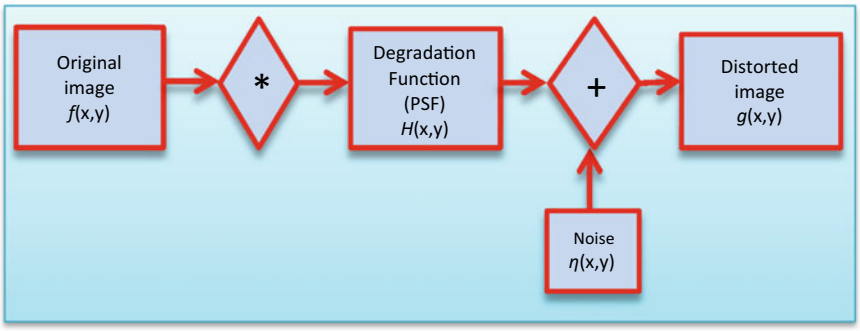


Fig. 3.4 Degradation process

3.3 Image Restoration Process

In this work, we suppose that the PSF is known. Our algorithm consists of two steps: Deconvolution in the Fourier domain and denoising in the wave atom domain.

The first step consists of regularized deconvolution in the Fourier space by the parametric Wiener filter whose purpose is to offset in part the PSF and reduce degradations. Let:

$$h(x,y)^{-1} * g(x,y) = \underbrace{f(x,y)}_{\text{original image}} + \underbrace{h(x,y)^{-1} * \eta(x,y)}_{\text{deconvoluted noise}} \tag{3.6}$$

where the term $h^{-1}(x,y) * \eta(x,y)$ is a term that corresponds to the noise. This deconvolution is provided by the parametric Wiener filter.

The second step consists in decomposing the resulting image, which still contains a deconvoluted noise part, into wave atom transform and reconstruct it after local thresholding of the coefficients.

The threshold choice is important since it will determine the level of detail to remember: a low threshold rebuilt the image with many details but also with a significant part of the noise. A high threshold outputs an image devoid of its fine details, and the noise is reduced. The algorithm steps are as follows:

Step 1. Distorted image using Eq. (3.3).
Step 2. Deconvolution in the Fourier domain using Eq. (3.6).
Step 3. Decomposition of the resulting image into wave atom domain.
Step 4. Setting local coefficients threshold and construction of the final image.

The restoration process is shown in Fig. 3.5.

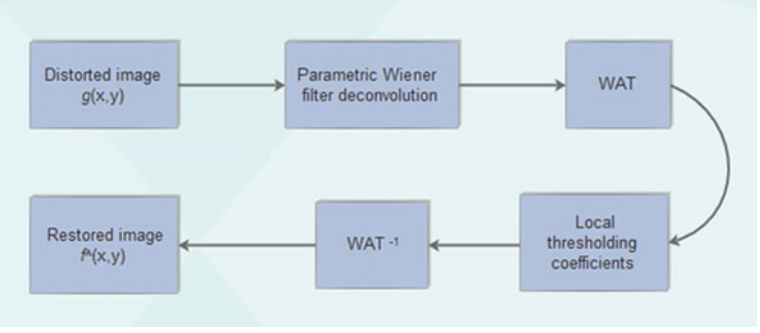


Fig. 3.5 Restoration process

3.3.1 Deconvolution in the Spectral Domain

Wiener filtering is one of the earliest and best known approaches to linear restoration. This filter seeks an estimate $\hat{f}$ that minimizes the statistical error function e^2.

$$e^2 = E\left[(f(x, y) - \hat{f}(x, y))^2\right] \tag{3.7}$$

where E is the expected value operator $f(x, y)$ is the original image and $\hat{f}(x, y)$ is the estimated image.

The Wiener deconvolution is a mathematical operation which involves applying a Wiener filter to eliminate or attenuate part of the noise in a signal. It operates in the frequency domain by trying to minimize the impact of noise where the signal-to-noise ratio is low. In this work, convolution is performed by the parametric Wiener filter which is a particular case of the wiener filter.

3.3.1.1 Wiener Estimator

Assuming first that $F(u, v)$ does not exist. The ideal image $f(x, y)$ and the noise $\eta(x, y)$ are stationary random functions and so are the degraded image $g(x, y)$ and the restored image. The mean square error $e^2(x, y)$ to be minimized is:

$$e^2(x, y) = \frac{1}{S} \iint_S e^2(x, y)\mathrm{d}x\mathrm{d}y \tag{3.8}$$

where $e(x, y) = f(x, y) - \hat{f}(x, y)$ and S is the image surface. The quadratic error scheme is shown in Fig. 3.6.

The impulse response $h_0(x,y)$ of the filter which leads to this result is called Wiener estimator as:

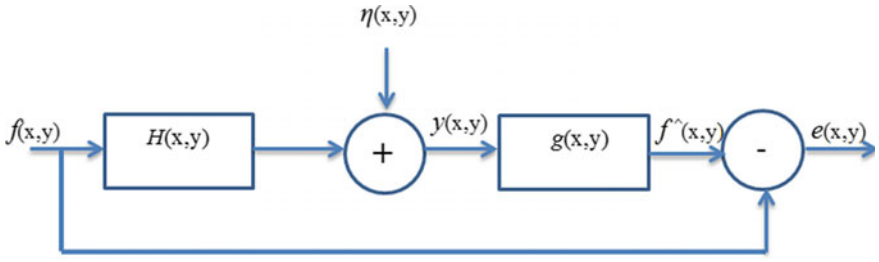


Fig. 3.6 Retro-propagation quadratic error decreasing

$$C_{\mathrm{gf}} = h_0 * C_g \tag{3.9}$$

where $C_g(x, y)$ is the autocorrelation function of the image $g(x, y)$ and $C_{\mathrm{gf}}(x, y)$ is the normalized cross-correlation function between images f and g.

The Wiener estimator transfer function deducted from the Fourier transform of Eq. 3.9 is:

$$H_0(u, v) = \frac{P_{\mathrm{gf}}(u, v)}{P_g(u, v)} \tag{3.10}$$

where P_{gf}, P_g are the Fourier transforms of C_{gf} and C_g, respectively.

If the signal and noise are uncorrelated, then.

$$P_g = P_f + P_\eta \quad \text{and} \quad P_{\mathrm{gf}} = P_f \tag{3.11}$$

where P_η (u, v) is the power spectrum of noise $\eta(x, y)$ and $P_f(u, v)$ is the power spectrum of the original image $f(x, y)$. So Eq. 3.10 becomes:

$$H_0(u, v) = \frac{P_f(u, v)}{P_f(u, v) + P_\eta(u, v)} \tag{3.12}$$

3.3.1.2 Wiener Filter Formulation

The Wiener deconvolution method has widespread use in image deconvolution applications, as the frequency spectrum of most visual images is fairly well behaved and may be estimated easily.

The goal of the Wiener filter is to compute a statistical estimate of unknown signal using a related signal as an input and filtering it to generate the estimate as output;

The Wiener filter transfer function is nothing other than the filter resulting from the cascading of the inverse filter and the Wiener estimator.

The solution to this model in the frequency domain is.

$$\hat{F}(u, v) = \left[\frac{1}{H(u, v)} \times \frac{|H(u, v)|^2}{|H(u, v)|^2 + S_\eta(u, v)/S_f(u, v)} \right] G(u, v) \tag{3.13}$$

where $H(u, v)$ is the degradation function, $|H(u, v)|^2$ the complex conjugate of $H(u, v)$, $S_\eta(u, v) = |N(u, v)|^2$ the power spectrum of the noise, and $S_f(u, v) = |F(u, v)|^2$ the power spectrum of the original image.

The ratio $S_\eta(u, v)/S_f(u, v)$ is called the noise-to-signal power ratio. We notice that if the noise power spectrum is zero for all relevant values of u and v, this ratio becomes zero and the Wiener filter is reduced to the inverse filter.

The two related quantities of interest are the average noise power and the average image power, defined as:

$$\eta_A = \frac{1}{\text{MN}} \sum_u \sum_v S_\eta(u, v) \tag{3.14}$$

and

$$f_A = \frac{1}{\text{MN}} \sum_u \sum_v S_f(u, v) \tag{3.15}$$

where as usual, M and N denote the number of rows and columns of the image and noise arrays, respectively. These quantities are scalar constants, and their ratio R is:

$$R = \frac{\eta_A}{f_A} \tag{3.16}$$

R, which is also a scalar, is used sometimes to generate a constant array instead of the function $S_\eta(u, v)/S_f(u, v)$. In this case, even if the actual ratio is not known, we need to experiment interactively by varying R and see the restored results. This, of course, is a crude approximation that assumes that the function is constant. Replacing $S_\eta(u, v)/S_f(u, v)$ by a constant array in the preceding filter equation results in the so-called parametric Wiener filtering. The deconvolution by the Wiener filter is shown schematically in Fig. 3.7.

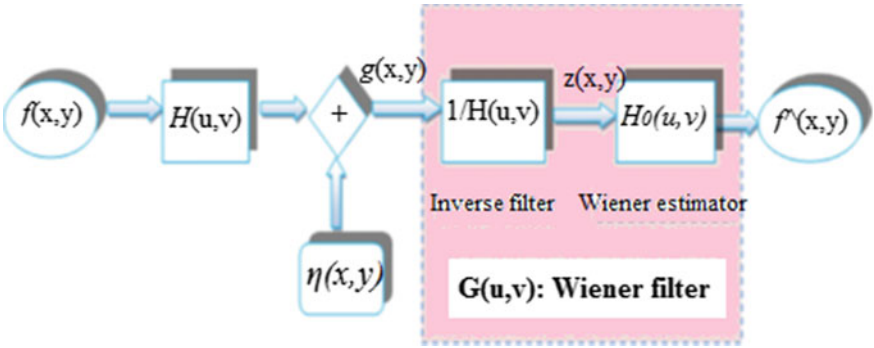


Fig. 3.7 Mathematical model of the wiener filter

3.3.2 Denoising in the Wave Atom Domain

3.3.2.1 Review of the Wave Atom Transform

During the last decade a series of variants of geometric transformations have appeared. In fact, traditional wavelet transformation has attracted much attention from researchers in the field of image analysis. But, in practice, this transformation has shown some limitations. Indeed, the transition from a level to another in conventional wavelet transform causes the approximation decomposition while in the wavelets packets, the decomposition could be pursued into the other sets (details and approximation), which is not optimal.

After a few years, Yong and Demanet presented a new transformation obeying the parabolic scaling law with a wavelength equal to diameter^2 [5, 12]. This transformation is considered a variant of 2D wavelet packets and an efficient representation of oscillating patterns and textures in a large class of images. The new transformation is called wave atom transform (WAT) and has two important proprieties:

- The ability to adapt to arbitrary local directions of a pattern.
- The ability to sparsely represent anisotropic patterns aligned with the axes.

The WAT has indicated a first convenience that can be adapted to arbitrary local directions of a second pattern and allows representing anisotropic models aligned with the axes. It also has another significant convenience that gives us a strong frequency location than other wave packets. Moreover, the oscillation functions or texture direction is sparser than other forms like wavelets transform and Gabor transform.

In image processing, the wavelets are used separably on the horizontal axis and the vertical axis which generates a partial decorrelation of the image giving many high energy coefficients along the contours, that is, why more adapted transforms to the representation of curves and outlines in an image, have been proposed. These transforms are known as geometric wavelet transforms.

The wave atoms transform is a geometric wavelet transform and created using two parameters, which are α and β with $0 \leq \alpha, \beta \leq 1$. These variables correspond successively to the multi-scale structure of the transform and directional ability. When $\alpha = 0$, the transform is not multi-scale such as the case of the Gabor transform. On the other hand, when $\alpha = 1$, the transform is multi-scale, this is the case of wavelets, ridgelets, and curvelets. Similarly, when $\beta = 0$, the transform has a maximum selectivity such as the Gabor transform where the decomposition can be done in all possible direction. However, $\beta = 1$ represents a minimal selectivity, and this is the case of wavelets transform where there is very little directional information. For the case of wave atoms, $\alpha = \beta = 1/2$. They interpolate precisely between Gabor atoms and directional wavelets [13, 14]. Figure 3.8 shows the multi-scale structure and directional ability for each transforms.

Wave atoms are constructed from tensor product of 1D dimensional wave packets and can be represented from wave packet 1D, $\psi^j_{m,n}(x)$ where $j, m > 0$ and $n \in Z$.

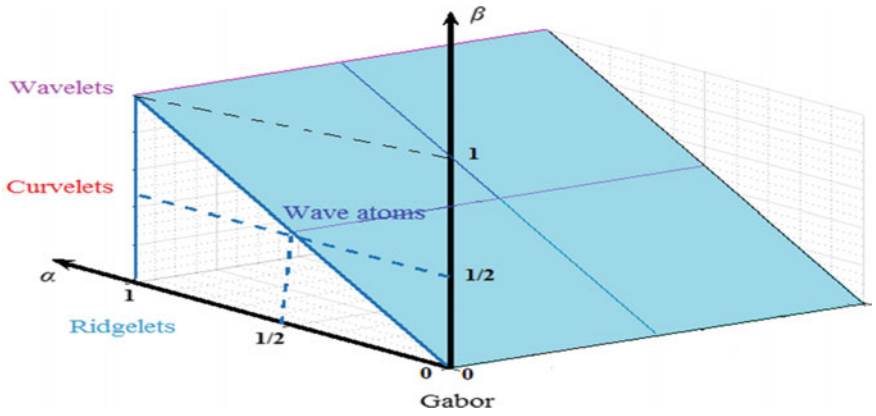


Fig. 3.8 Multi-scale structure and directional ability for different transforms

The wave packets 1D are centered around the $X_{j,n} = 2^{-j}n$and $\mp\psi_{j,m} = \mp\pi 2^j m$ in space and frequency domains, respectively, with $C_1 2^j < m < C_2 2^j$.

The basic function in frequency domain is obtained from a combination between dyadic scaled and translated version of $\hat{\psi}_m^0$ and can represented by Eq. (3.17):

$$\psi_{m,n}^j(x) = \psi_m^j(x - 2^{-j}n) = 2^{j/2}\psi_m^0(2^j x - n) \tag{3.17}$$

where $\psi_m^0(\omega)$ is defined as:

$$\psi_m^0(\omega) = e^{-i\omega/2}\left[e^{i\alpha_m} g(\ell_m(\omega - \pi(m + 1/2))) + e^{-i\alpha_m} g(\ell_{m+1}(\omega + \pi(m + 1/2)))\right] \tag{3.18}$$

With $\alpha_m = \pi/2(m + 1/2)$, $\ell_m = (-1)^m$ and g, a real-value compactly supports C^∞ bump function such that $\sum_m \left|\psi_m^0(\omega)\right|^2 = 1$.

For each wave $\omega_{j,m}$ at scale 2^{-j}, the coefficient $c_{j,m,n}$ can be treated as a decimated convolution as follows:

$$c_{j,m,n} = \int \psi_m^j(x - 2^{-j}n)u(x)\mathrm{d}x = 1/2\pi \int e^{i2^{-jn\omega}} \overline{\hat{\psi}_m^j(\omega)}\hat{u}(\omega)\mathrm{d}\omega \tag{3.19}$$

By discretizing the sample u at $x_k = \mathrm{kh}$, $h = 1/N$, $k = 1 \ldots N$, the discrete coefficients $c_{j,m,n}^D$ are computed using a reduced inverse *FFT* inside an interval of size $2^{j+1}\pi$, centered around the origin.

$$c_{j,m,n}^D = \sum_{k=2\pi(-2^j/2+1:1:2^j/2)} e^{i2^{-j}nk} \sum_{p\in 2\pi Z} \overline{\hat{\psi}_m^j(k + 2^j p)}\hat{u}(k + 2^j p) \tag{3.20}$$

Introducing subscript $\mu = (j, m_1, m_2, n_1, n_2)$, 2D orthonormal basis is given by Eq. 3.21 and 3.22:

$$\varphi_{\mu}^{+}(x_1, x_2) = \psi_{m1}^{j}(x_1 - 2^{-j}n_1) \quad \psi_{m2}^{j}(x_2 - 2^{-j}n_2) \tag{3.21}$$

and

$$\varphi_{\mu}^{-}(x_1, x_2) = H\psi_{m1}^{j}(x_1 - 2^{-j}n_1)\ H\psi_{m2}^{j}(x_2 - 2^{-j}n_2) \tag{3.22}$$

where H is a Hilbert transform.

Combining Eqs. (3.21) and (3.22) forms the wave atom tight frame as follows:

$$\varphi_{\mu}^{(1)} = \frac{\varphi_{\mu}^{+} + \varphi_{\mu}^{-}}{2} \tag{3.23}$$

and

$$\varphi_{\mu}^{(2)} = \frac{\varphi_{\mu}^{+} - \varphi_{\mu}^{-}}{2} \tag{3.24}$$

Figures 3.9 and 3.10 show, respectively, an example of wave atom 1D and 2D in spatial domain and Fourier domain.

3.3.2.2 Nonlinear Threshold Optimization

The threshold choice is important because it will determine the level of details to remember. A low threshold rate gives an image with great detail but also with a significant part of the noise. On the other hand, a high threshold produces an image without fine details, and the noise is reduced.

There are several types of threshold estimators. We can first distinguish hard thresholding and soft thresholding.

A. ***Hard thresholding***

Hard thresholding is the one that is the most "intuitive." We set a threshold $T > 0$. We only keep the wavelet coefficients greater than T, and we set the others to zero. The shape of hard threshold is shown in Fig. 3.11.

$$T(x) = \begin{cases} 0 \text{ if } |x| < T \\ x \text{ if } |x| \geq T \end{cases} \tag{3.25}$$

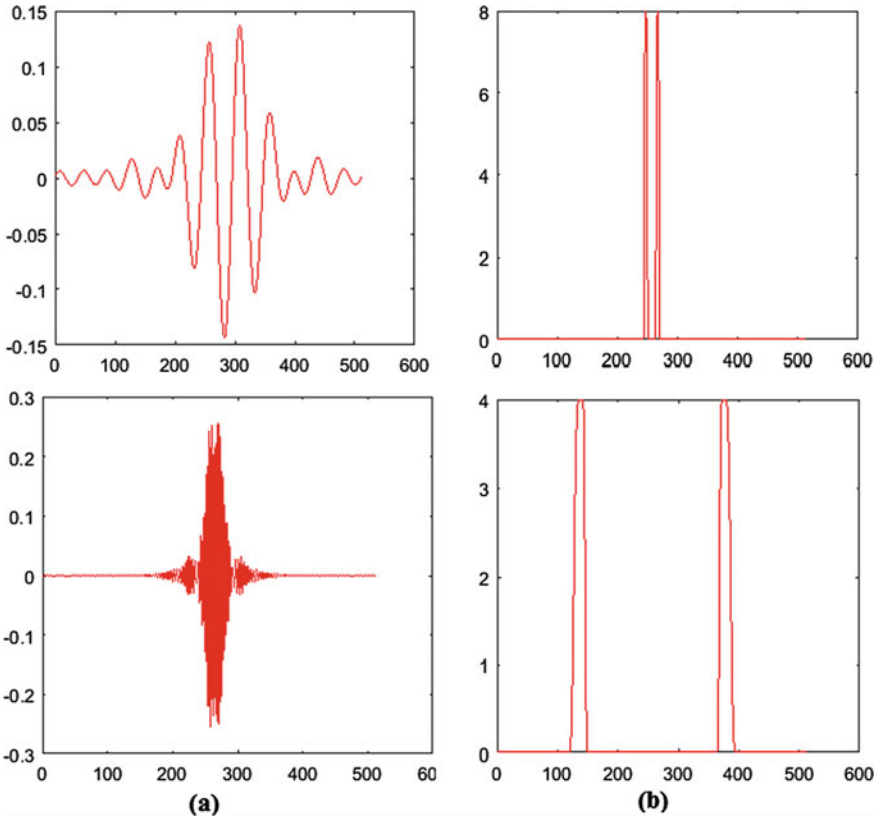


Fig. 3.9 1D wave atom transform ($j = 3$, $m = 3$; $j = 5$, $m = 8$), **a** spatial domain, **b** frequency domain

B. ***Soft thresholding***

In the case of soft thresholding, the coefficients lower than a threshold T are always set to zero

On the other hand, for those higher than T, the amplitude of the coefficients is attenuated by the value of the threshold in order to be sure to have the effect of the noise removed even for the high coefficients. Figure 3.12 shows the shape of the soft thresholding.

$$T(x) = \begin{cases} 0 & \text{if } |x| < T \\ x - \text{sign}(x) \times T & \text{if } |x| \geq T \end{cases} \tag{3.26}$$

Both thresholding functions above have advantages and disadvantages. In fact, the soft thresholding is not effective for large coefficients. On the other hand, for hard

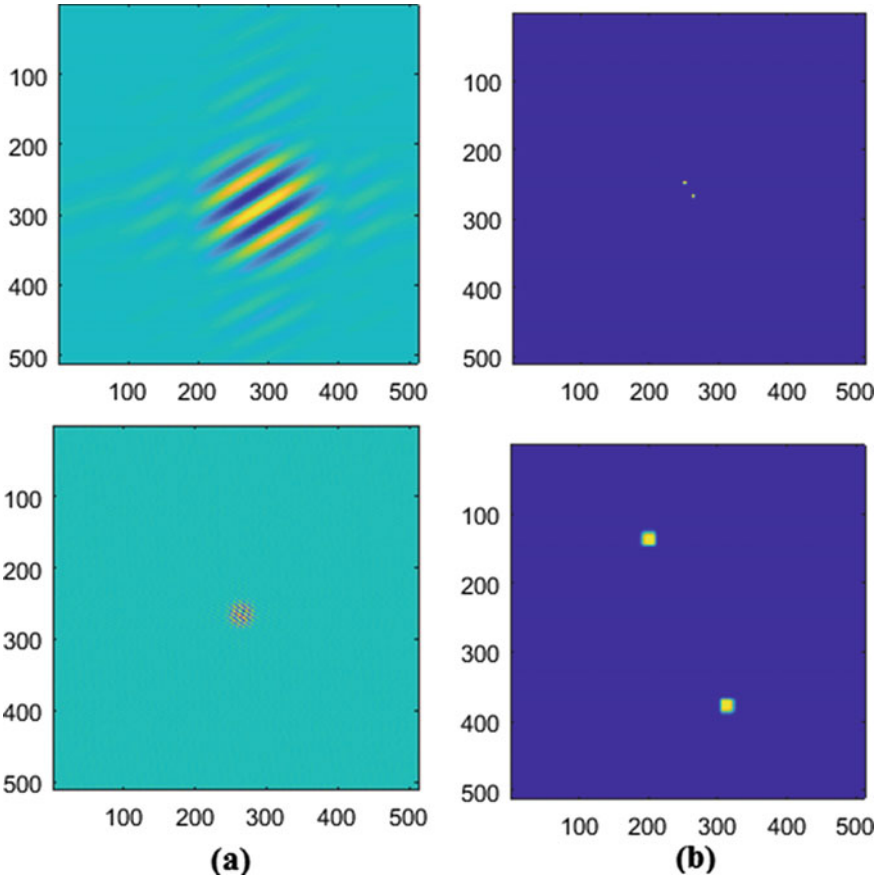


Fig. 3.10 2D wave atom transform ($j = 3$, $m = 3{,}2$; $j = 5$, $m = 8{,}4$), **a** spatial domain, **b** frequency domain

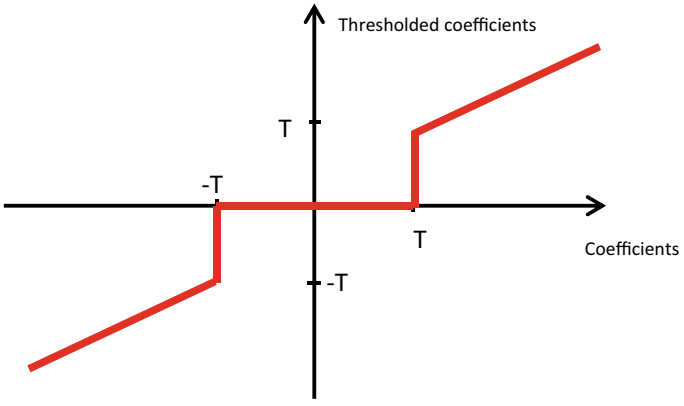


Fig. 3.11 Hard threshold

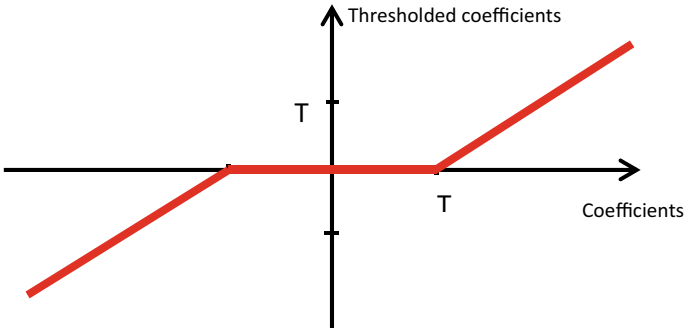


Fig. 3.12 Soft threshold

thresholding, the results tend to have a large variance and to be unstable because of the discontinuities of the function of thresholding.

In the literature, there are several methods of thresholding, and the most known and applied are:

– The VisuShrink threshold: Introduced by Donoho and Johnstone [15, 16], also called universal threshold. This method uses the assumption of Gaussian white noise superimposed on the signal. The threshold is equal to.

$$S = \sigma_n \sqrt{2 \log_e N} \tag{3.27}$$

where σ_n is the noise variance and N is the image size.

Since the noise is considered white, its energy is equi-distributed over all the frequency bands of the decomposition. Accordingly, the standard deviation is estimated in a band where the informational signal is considered non-existent or almost non-existent, particularly in the higher frequencies. In their work, the authors use a robust estimator of σ_n from the median value of the coefficients of the diagonal detail sub-band of the first decomposition level.

$$\sigma_n^2 = \left[\frac{\text{median}(|y_{ij}|)}{0.6745} \right], y_{\text{ij}} \in \quad \text{subband HH}_1 \tag{3.28}$$

This threshold is effective if the informative signal is hollow, i.e. if its coefficients are rare. In other words, the thresholding may be too strong for any signal. Furthermore, universal thresholding has another problem which is the unique value of the threshold for all frequency bands of decomposition.

– The BayesShrink threshold: It uses a Bayesian mathematical structure for the images to obtain specific thresholds for each sub-band [17]. For each sub-band, the threshold is given by:

$$T_B = \frac{\hat{\sigma}_n^2}{\hat{\sigma}_x} \tag{3.29}$$

where $\hat{\sigma}_n^2$ is the estimated noise variance and $\hat{\sigma}_x$ is the estimated standard deviation of the signal.

$$\hat{\sigma}_x = \sqrt{\max(\hat{\sigma}_y^2 - \hat{\sigma}_n^2, 0)} \tag{3.30}$$

With $\hat{\sigma}_y^2$ is the estimated variance of the observed signal which is written.

$$\hat{\sigma}_y^2 = \frac{1}{M^2} \sum_{i,j=1}^{M} Y_{ij}^2 \tag{3.31}$$

– The Normal shrink threshold: It is an adaptive threshold since the parameters required to evaluate the effectiveness of the threshold depend on the data of the sub-bands [18]. This threshold is then calculated by:

$$T_N = \beta \hat{\sigma}_n^2 / \hat{\sigma}_x \tag{3.32}$$

where the $\hat{\sigma}_n^2$ is the estimated noise variance, the $\hat{\sigma}_x$ is the estimated standard deviation of the signal, and β is the scale parameter, which depends on the size of the sub-band and the level of decompositions given by Eq. 3.33:

$$\beta = \sqrt{\log(L_k / J)} \tag{3.33}$$

where L_k is the length of the sub-band at the kth scale.

In this work, we suggest to use the VisuShrink threshold with the estimation the noise variance. The noise parameters are typically estimated by analyzing the Fourier spectrum. Let f_i denotes intensity levels in an image and $P(f_i)$, i = 0, 1, 2,. . .L−1, be the corresponding normalized histogram, where L is the number of possible intensity values. A histogram component, $P(f_i)$, is an estimate of the probability of occurrence of intensity value f_i and the histogram may be viewed as a discrete approximation of the intensity of probability density function (PDF).

One of the principal approaches for describing the shape of a histogram is the use of its central moments (also called moments about the mean), which are defined as follows:

$$\mu_n = \sum_{i=0}^{L-1} (f_i - m)^n P(f_i) \tag{3.34}$$

where n is the moment order and m the mean defined as follows.

$$m = \sum_{i=0}^{L-1} f_i P(f_i) \tag{3.35}$$

Because the histogram is assumed to be normalized, the sum of all its components is equal to 1. So, we can see from the preceding equation that $\mu_0 = 1$ and $\mu_1 = 0$, and the second-order moment is the variance.

By selecting an area of a nearly constant background level, and assuming that the noise is additive, we can estimate that the average intensity of the area in the ROI is reasonably close to the average gray level of the image in that area without noise. This indicates that the noise in this case has zero mean. Also the fact that the area has a nearly constant intensity level tells us that the variability in the region in the ROI is primarily due to the variance of the noise.

To estimate the variance of the noise, first the image is segmented using the watershed algorithm, as it is shown in Fig. 3.13, and then, the second-order moment is calculated for each region.

After the estimate of the noise variance, the threshold is calculated on windows of size 64 × 64 based on the following principle:

- If the window is homogeneous, the threshold is calculated according to the noise variance in this zone.
- If the window contains multiple areas, the threshold is calculated as the average of the estimated variances in these areas.

After the variance is estimated, each window is filtered by using the corresponding threshold, which is calculated using the estimated variance. The optimal threshold is given in Fig. 3.14.

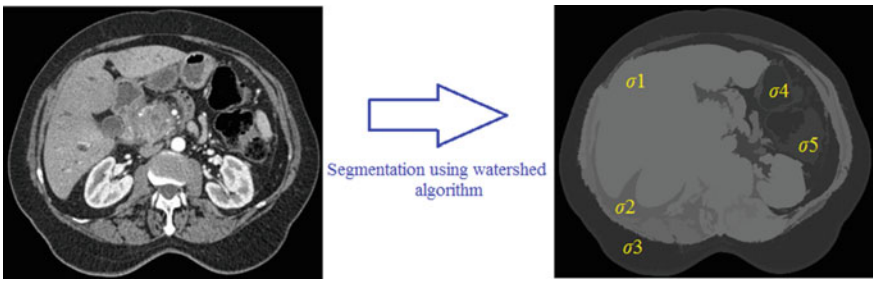


Fig. 3.13 Noise variance estimation

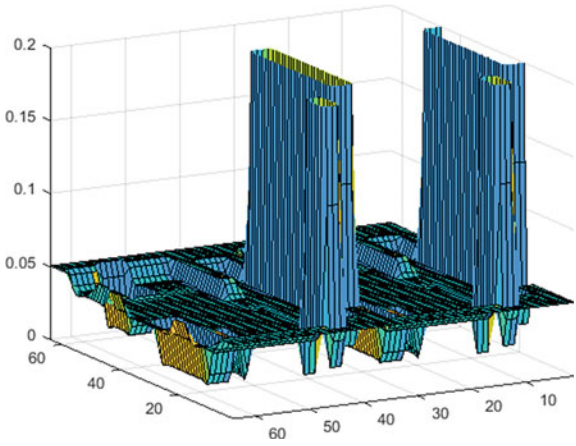


Fig. 3.14 Adaptive nonlinear threshold calculated with overlapped windows size (8 × 8)

3.4 Experimental Results

In order to test the effectiveness of our algorithm objectively and subjectively, experiments were performed on simulated low resolution images. Those images were initially degraded by variable spread functions (Gaussian blur, motion blur). In a second step, the obtained images were further degraded by adding an additive Gaussian noise. The resulting images were then restored by three methods: the classical Wiener method, deconvolution into wavelet domain method, and the proposed method. Different measures were calculated with the aim to compare the results of filtering. In fact, we used a set of criteria. The first criterion that we used is the PSNR, given by Eq. 3.36.

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{d^2}{\mathrm{MSE}}\right) \tag{3.36}$$

With:

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left([f(x,y) - r(x,y)]^2\right) \tag{3.37}$$

The second criterion is the normalized cross-correlation (NCC). The third one is the index of structural similarity, known by structural similarity index (SSIM) given in Eq. 3.38, and the fourth criterion is the edge preserving index given in Eq. 3.39.

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3.38}$$

With μ_x is the x average, μ_y the y average, σ_x^2 the x variance, σ_y^2 the y variance, σ_{xy} the x and y covariance, $c_1=(k_1\ L)^2$ and $c_2=(k_2\ L)^2$ two variables to stabilize the division with weak denominator and L the dynamic range of the pixel-values. K_1=0.01 and k_2=0.03 by default.

$$\mathrm{EPI} = \frac{\sum_{i,j} |f_s(i,j) - f_s(i-1, i+1)|}{\sum_{i,j} |f_i(i,j) - f_i(i-1, i+1)|} \tag{3.39}$$

where $f_s(i, j)$ is the value of the smoothed image pixel and $f_i(i, j)$ is the value of the original image pixel. The index i is the row number and j is the column number.

The last criterion is based on the edge detection of restored images using the canny operator with threshold equal to 0.1.

3.4.1 Numerical Results Using Gaussian Blur

Many experiments are recorded in order to show the important contribution of the new algorithm. Images are blurred by Gaussian blur with size of 21 × 21, then an additive zero mean Gaussian noise is added with different values of BSNR. The results obtained are shown in Figs. 3.15, 3.16, 3.17, 3.18, and 3.19 and Tables 3.1 and 3.2.

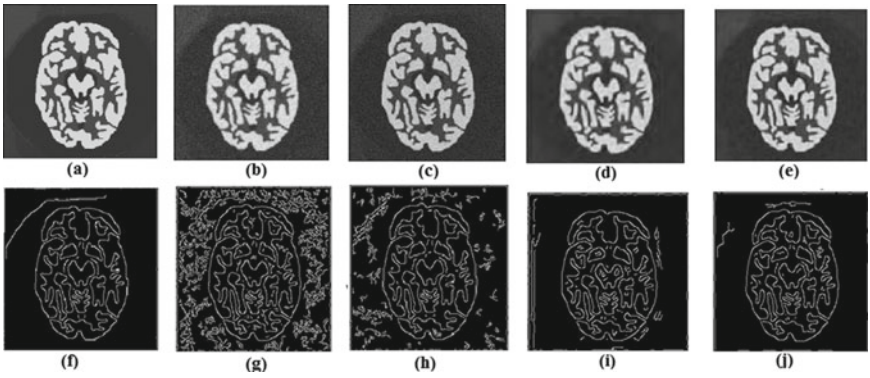


Fig. 3.15 MRI image, **a** original image, **b** degraded image (Gaussian blur, BSNR = 10 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

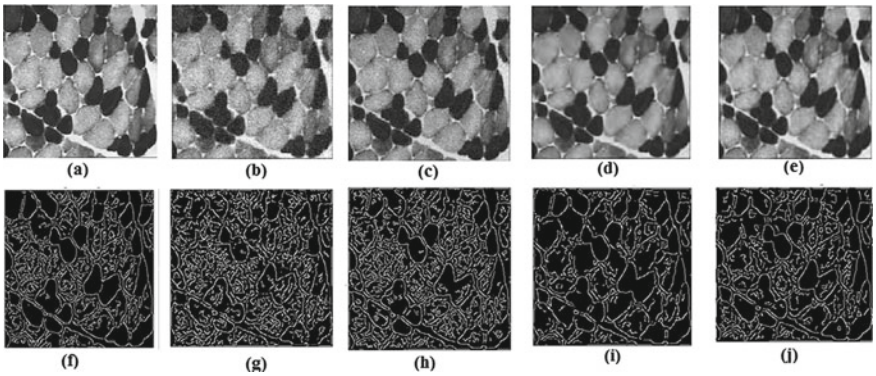


Fig. 3.16 Cell image, **a** original image, **b** degraded image (Gaussian blur, BSNR = 15 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

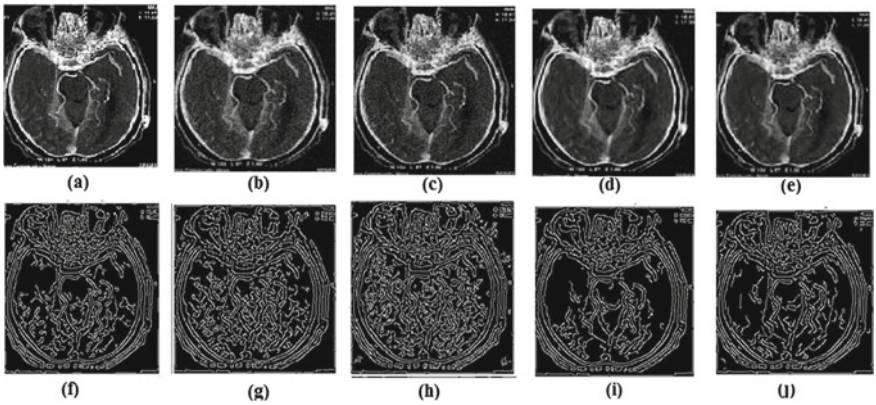


Fig. 3.17 Stomach image, **a** original image, **b** degraded image (Gaussian blur, BSNR = 15 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

3.4.2 Numerical Results Using Motion Blur

To test the effectiveness of the proposed algorithm against the motion blur, images are blurred by motion blur kernel with the angle parameter ($\theta = 135°$) and the distance parameter (Len = 11) and corrupted by additive zero mean Gaussian noise with BSNR = 20 dB. The results obtained are compared with those obtained by classic wiener filtering and wavelet method. Simulation results are given in Figs. 3.20, 3.21, 3.22, 3.23, and 3.24 and Tables 3.3 and 3.4.

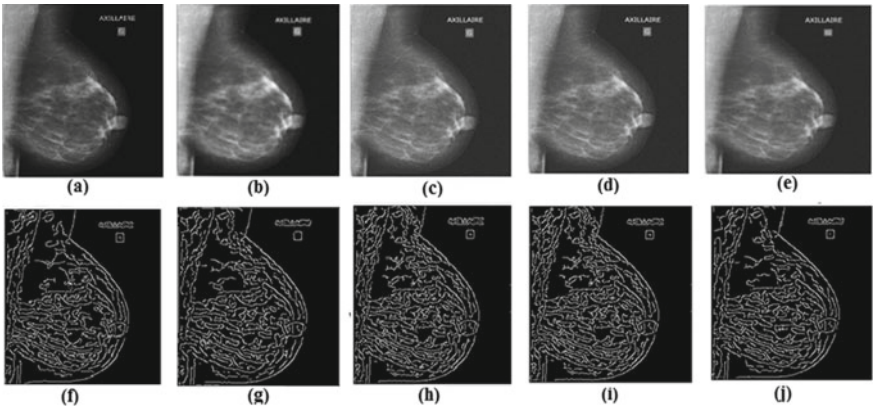


Fig. 3.18 Breast image, **a** original image, **b** degraded image (Gaussian blur, BSNR = 20 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

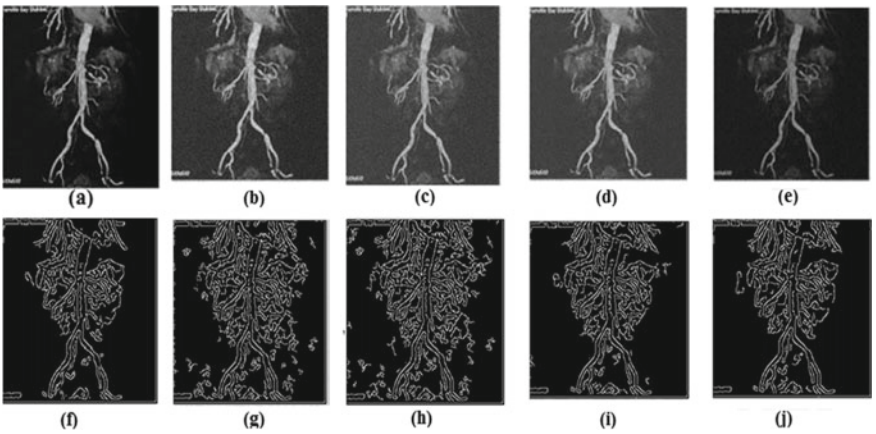


Fig. 3.19 Stenosis image, **a** original image, **b** degraded image (Gaussian blur, BSNR = 15 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

3.5 Discussion

From the obtained results, we can show that the proposed method performed other methods in terms of PSNR, cross-correlation SSIM, and EPI in the majority of cases. In fact, the gap between the PSNR obtained by the proposed method and the classical Wiener filter is almost about 3 dB.

These results were confirmed by visual examination. Furthermore, by using wiener filtering, the blur has been significantly reduced. In contrast, in the case

Table 3.1 Numerical results using Gaussian blur

	Wiener filtering method					Wavelet method				Proposed method			
Image	R	PSNR	NCC	SSIM	EPI	PSNR	NCC	SSIM	EPI	PSNR	NCC	SSIM	EPI
MRI	0.04	22.2709	0.9799	0.9799	0.4684	24.6874	0.9763	0.9799	0.6207	**25.3377**	**0.9799**	**0.9999**	**0.8254**
Cell	0.05	22.4197	0.9698	0.9996	0.9950	22.4616	0.9726	0.9996	1.5370	**23.8184**	**0.9792**	**0.9998**	**1.8429**
Stomach	0.03	20.7365	0.9449	0.9997	1.0394	20.8822	0.9507	0.9997	1.3226	**21.3263**	**0.9526**	**0.9998**	**1.6683**
Breast	0.01	19 .3358	0.9801	0.5963	0.3539	19 .8378	0.9816	0.5980	0.3845	**21.1263**	**0.9936**	**0.6208**	**0.8534**
Stenosis	0.01	19.8066	0.8780	0.9998	0.1779	21.5711	0.9103	0.9999	0.2203	**24.2196**	**0.9914**	**0.9996**	**1.2194**

Table 3.2 Computing time in (s) for different methods (Gaussian blur)

Image	Wiener filtering method	Wavelet method	Proposed method
MRI	02.44	02.93	04.96
Cell	02.69	02.80	04.51
Stomach	02.10	02.51	05.13
Breast	02.35	02.70	05.21
Stenosis	02.83	02.96	05.33

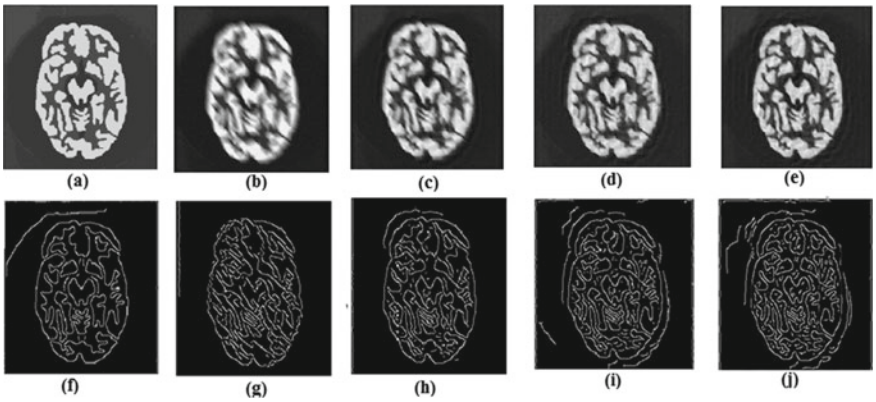


Fig. 3.20 MRI image, **a** original image, **b** degraded image (motion blur, BSNR = 20 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

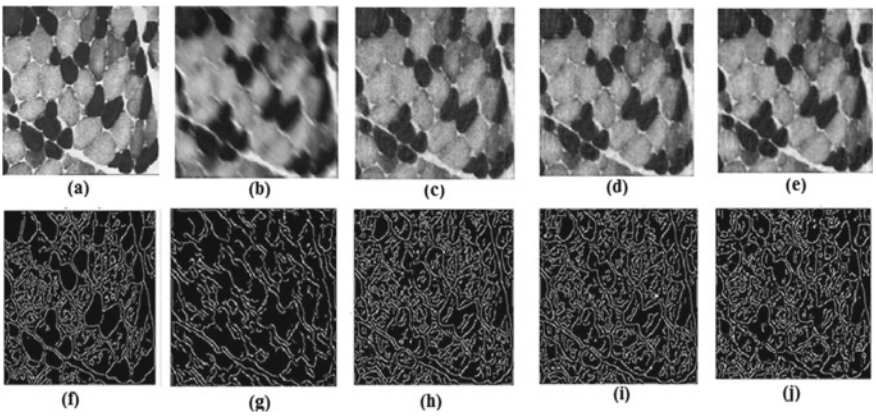


Fig. 3.21 Cell image, **a** original image, **b** degraded image (motion blur, BSNR = 20 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

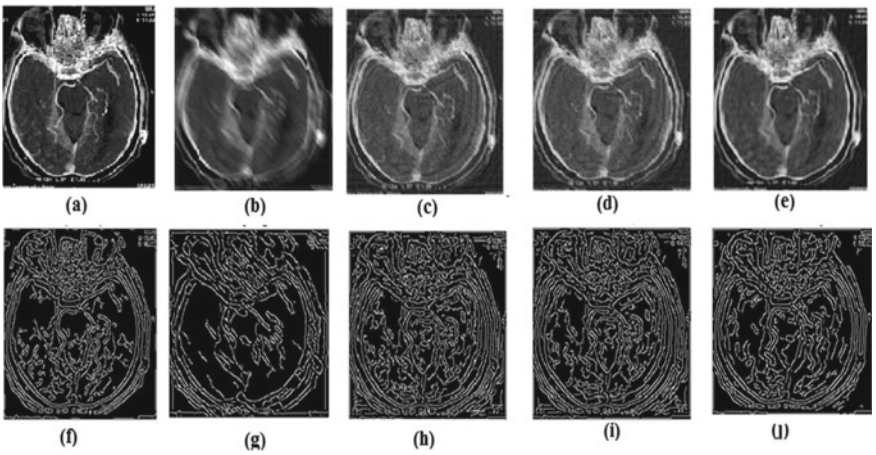


Fig. 3.22 Stomach image, **a** original image, **b** degraded image (motion blur, BSNR = 20 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

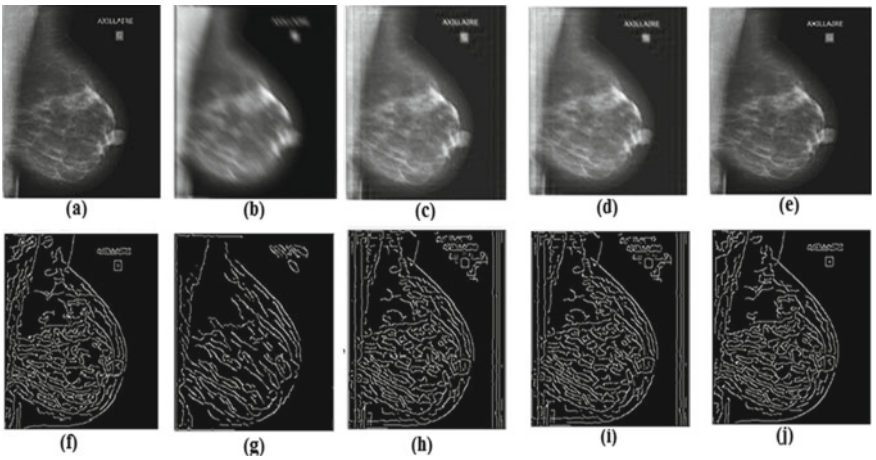


Fig. 3.23 Breast image, **a** original image, **b** degraded image (motion blur, BSNR = 20 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

of wavelet method, images were smoothed, and the noise is slightly removed. This explains the ability of wavelets method to compress the image in a few nonzero coefficients, which facilitates the denoising operation. On the other hand, the homogeneous regions in the images obtained by the proposed method have been smoothed, and the shapes and textures of the original images were found. This shows that the proposed method is a trade-off between noise removing and edge preserving.

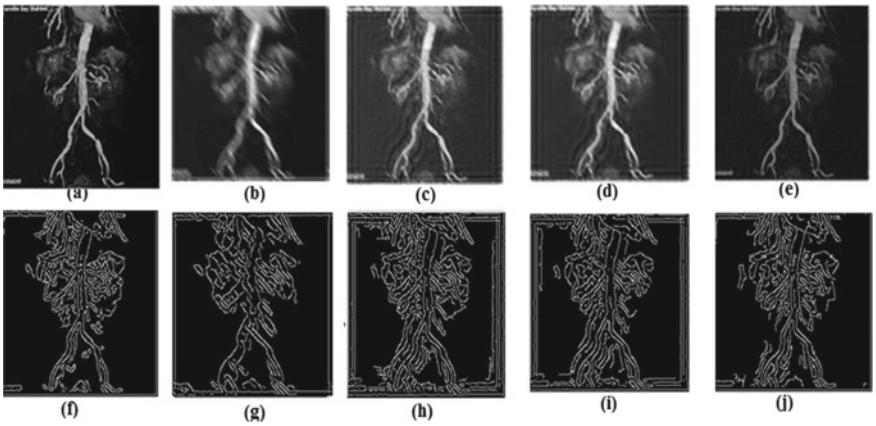


Fig. 3.24 Stenosis image, **a** original image, **b** degraded image (motion blur, BSNR = 20 dB), **c** Wiener filter, **d** wavelet filter, **e** proposed method, **f** original edge, **g** noisy edge, **h** Wiener edge, **i** wavelet edge, **j** proposed method edge

To confirm these findings, we propose to detect edges in the different images using a simple derivative operator. This will test the skill of methods to preserve the transitions. In this regard, the results show that the proposed method marks its superiority in terms of preserving contours. In fact, the Wiener method retained some of the noise which disrupted the edge detector and generated a false edge. The deconvolution by wavelets method retained the image transitions, and sometimes it generated open contours. The edge obtained by the proposed method, however, was clear closed and continuous, and details were well-preserved.

For the real-time aspect, we calculated the computing time using a CPU with 2.30 GHz as frequency. The results given in Tables 3.2 and 3.4 show that the necessary time of the proposed method is superior to the necessary time of the classical wiener filter and wavelet method. This is reasonable, since the proposed method requires an additional time to calculate the threshold for each area of the image.

3.6 Conclusion

In this work, we presented a rapid hybrid algorithm for supervised medical image restoration that combines both parametric Wiener filter and wave atoms transform. The importance of our method is the ability to restore the image without a priori knowledge of the spectral density of the original image and the noise. The proposed algorithm consists of two steps: deconvolution in the Fourier domain using parametric wiener filtering followed by denoising in the wave atom domain after local thresholding of the coefficients.

Table 3.3 Numerical results using motion blur

Image	Wiener filtering method					Wavelet method				Proposed method			
	R	PSNR	NCC	SSIM	EPI	PSNR	NCC	SSIM	EPI	PSNR	NCC	SSIM	EPI
MRI	0.01	22.0797	0.9510	0.9994	0.5048	23.3644	0.9645	0.9997	0.7730	**24.4950**	**0.9712**	**0.9998**	**0.8203**
Cell	0.09	23.5935	0.9651	0.9908	0.8354	24.7856	0.9667	0.9998	0.9062	**25.1243**	**0.9690**	**0.9999**	**1.4955**
Stomach	0.03	20.8645	0.9386	0.9997	0.9563	20.9066	0.9377	0.9997	1.0994	**22.3132**	**0.9562**	**0.9999**	**1.4754**
Breast	0.01	17.6181	0.9868	0.8899	1.2609	18.8789	0.9827	0.8778	1.4118	**20.2723**	0.9989	0.9932	1.9218
Stenosis	0.02	23.5913	0.9374	0.9999	0.4940	23.3529	0.9372	0.9998	1.3709	**24.5719**	0.9796	0.9999	1.4585

Table 3.4 Computing time in (s) for different methods (motion blur)

Image	Wiener filtering method	Wavelet method	Proposed method
MRI	02.31	03.10	05.20
Cell	02.83	02.95	06.21
Stomach	02.76	02.91	06.13
Breast	03.25	03.40	06.21
Stenosis	03.01	03.06	05.75

To test the effectiveness of our algorithm, the proposed method was compared to the Wiener method and the wavelets method. Experiment results are very interesting and show the efficiency of our algorithm compared to other algorithms. In fact, the proposed method can improve others methods and shows its superiority in terms of PSNR, SSIM, NCC, and EPI. These finding is confirmed by visual examination and prove that the proposed method constitutes a trade-off between noise removing and edge preserving.

As prospects, our results can be extended to unsupervised restoration and can be applied to color images pertaining to different areas like archeology, color photography, and old handwritten documents.

Acknowledgements The author is grateful to the anonymous reviewers and the associate editor for their valuable comments, which have greatly helped to improve this work.

References

1. Yang L, Zhang X, Ren J (2011) "Adaptive wiener filtering with Gaussian fitted point spread function in image restoration." *Software engineering and service science* (*ICSESS*), *IEEE 2nd international conference*, Beijing, pp 890–894
2. Qing Qin F (2012) "Blind image restoration based on Wiener filtering and defocus point spread function estimation, " In 5th international congress on image and signal processing (CISP 2012), Chongqing, Oct 2012, pp 360–363
3. Dai W, Ye Y (2007) "Image de-noising based on combination of Wiener filter and wavelet shrinkage", In Proceedings of the 2007 IEEE international conference on integration technology, Mar 20–24, Shenzhen, China
4. Li X, Su X, Ji L (2010) "Image de-noising via doubly Wiener filtering with adaptive directional windows and mean shift algorithm in wavelet domain, " In Proceedings of the 2010 IEEE international conference on mechatronics and automation, Aug 4–7, 2010, Xi'an, China
5. Donoho D, Johnstone I (1994) "Ideal spatial adaptation via wavelet shrinkage," Biometrika 81(3):425–455
6. Afonso MV, Bioucas-Dias JM, Figueiredo MAT (2010) Fast image recovery using variable splitting and constrained optimization. IEEE Trans Image Process 19(9):2345–2356
7. Pham TD (2012) Supervised restoration of degraded medical images using multiple-point geostatistics. Comput Methods Programs Biomed 06:201–209
8. Li D, Tian X, Jin Q, Hirasawa K "Adaptive fractional-order total variation image restoration with split Bregman iteration" ISA Transactions Available online 12 Oct 2017

9. Tang S, Gong W, Li W, Wang W (2014) Non-blind image deblurring method by local and nonlocal total variation models. Sig Process 94:339–349
10. Yan C, Zhang Y, Jizheng X et al (2014) A highly parallel framework for HEVC coding unit partitioning tree decision on many-core processors. IEEE Sig Process Lett 21(5)
11. Yan C, Zhang Y, Xu J et al (2014) Efficient parallel framework for HEVC motion estimation on many-core processors. IEEE Trans Circ Syst Video Technol, http://dx.doi.org/10.1109/TCSVT.2014.2335852
12. Palakkal S, Prabhu KMM (2012) Poisson image denoising using fast discrete curvelet transform and wave atom. Signal Process 92:2002–2017
13. Demanet L, Ying L (2007) Wave atoms and sparsity of oscillatory patterns. Appl Comput Harmon Anal 23:368–387
14. Liu F, Cheng LM, Leung HY, Qi-Kai F (2012) Wave atom transform generated strong image hashing scheme. Opt Commun 285:5008–5018
15. Naimi H, Houda Adamou-Mitiche AB, Mitiche L (2015) Medical image denoising using dual tree complex thresholding wavelet transform and Wiener filter. J King Saud Univ Comput Inform Sci 27(1):40–45
16. Donoho D (1995) "De-noising by soft-thresholding". IEEE Trans Inform Theory 41(3):613–627
17. Chang S, Yu B, Vetterli M (2000) Adaptive wavelet thresholding for image denoising and compression. IEEE Trans Image Process 9(9):1532–1546
18. Kaur L, Gupta S, Chauhan RC (Dec 16–18 2002) "Image denoising using wavelet thresholding", ICVGIP, proceeding of the third indian conference on computer vision, graphics & image processing, Ahmdabad, India

Part II
Security and Privacy in IoT-Based E-Health Applications

Chapter 4
BlockCare: SDN-Enabled Blockchain Framework for Securing Decentralized Healthcare and Precision Medicine Applications

Akram Hakiri, Aniruddha Gokhale, and Nicolae Tapus

Abstract The growing importance and maturity of Internet of Things (IoT) and wearable computing are revolutionizing healthcare diagnosis and body treatment by providing access to meaningful healthcare data and improving the effectiveness of medical services. In this context, personal health information must be exchanged via trusted transactions that provide secure and encrypted sensitive data of the patient. Moreover, healthcare smart devices need flexible, programmable, and agile networks to allow on-demand configuration and management to enable scalable and interoperable healthcare applications. Two complementary trends show promise in meeting these needs. First, blockchain is emerging as a transparent, immutable, and validated-by-design technology that offers a potential solution to address the key security challenges in healthcare domains by providing secure and pseudo-anonymous transactions in a fully distributed and decentralized manner. Second, software-defined networking (SDN) offers a significant promise in meeting the healthcare communication needs by providing a flexible and programmable environment to support customized security policies and services in a dynamic, software-based fashion. To that end, we present our ideas on SDN-enabled blockchains that can be used to develop and deploy privacy-preserving healthcare applications. First, we present a survey of the emerging trends and prospects, followed by an in-depth discussion of major challenges in this area. Second, we introduce a fog computing architecture that interconnects various IoT elements, SDN networking, and blockchain computing components that control and manage patients' health-related parameters. Third, we

A. Hakiri (✉)
Department of Computer Science and Telecoms, University of Carthage, SYSCOM-ENIT, ISSAT, 7030 Mateur, Tunisia
e-mail: akram.hakiri@issatm.rnu.tn

A. Gokhale
Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235, USA
e-mail: a.gokhale@vanderbilt.edu

N. Tapus
Faculty of Automatic Control and Computers, University Politehnica of Bucharest, 060042 Bucharest, Romania
e-mail: nicolae.tapus@cs.pub.ro

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_4

validate our architecture in the context of three use cases involving smart health care, precision medicine, and pharmaceutical supply chain. Finally, we discuss open issues that need significant new research investigations.

4.1 Introduction

Recent advances in Internet of Things (IoT) smart devices for health care have enabled patient-to-doctor (P2D), patient-to-machine (P2M), and doctor-to-machine (D2M) connectivity [15]. Many healthcare providers adopt IoT solutions and wellness applications to help patients follow their prescribed care programs and facilitate analytics. The patients can gradually become digital stewards of their health data by using healthcare applications to monitor their conditions at home and communicate the state of their health with their family and doctors. However, sharing such confidential patient's data has placed an increasing strain on the security of the network infrastructure because healthcare systems can encounter severe security and privacy concerns, e.g., if unauthorized access is granted to wearable gadgets [2], or trustless transactions are performed between healthcare smart devices themselves and with their healthcare service providers over the network. All this has motivated the need for providing trusted transactions. Additionally, a compromised IoT healthcare device could be prone to distributed denial-of-service (DDoS) attacks and overwhelm hospital networks with malicious traffic that will make their services unavailable [28].

Beyond the issues of securing the transactions involving patient health records, challenges in securing the network still persist. For example, consider telemedicine practice where the patient's data are exchanged securely and scalably to enable clinical communication with patients at remote locations. The data exchange helps to gather recommendations and confirmation from a group of medical surgeons, consultants, and physicians with different specialties. These healthcare experts can meet virtually to analyze patient cases, collaborate to create effective treatment, and prevent inadequacies and errors in care plans and medication. Such virtual medical interactions could increase the risk of breach of clinical data due to transmitting data over untrusted networks thereby violating data privacy and revealing sensitive medical identity. Therefore, a trusted network infrastructure and trusted relationship between healthcare entities are required to ensure tamper-proof data sharing among healthcare stakeholders (e.g., care providers, healthcare institutions, insurance companies, etc.) over the network.

Furthermore, the healthcare industry is under pressure from both the perspective of keeping costs low and dealing with government regulations [14]. On the one hand, as the healthcare industry attempts to incorporate disruptive innovations, it becomes hard for healthcare stakeholders to remain competitive and keep their costs down. On the other hand, the lack of interoperability between healthcare providers and hospital systems forces governments and federal regulators to expect interoperable data standards and demand secured health information exchange (HIE) and electronic health record (EHR) systems. As a result, many healthcare providers err on the side of

caution when interpreting security standards or general requirements, e.g., HIPAA compliance, fast healthcare interoperability resources (FHIR), etc., for protecting health information [21]. This imposes too many restrictions on patients and their proxies from accessing data about their own health.

To address these challenging issues, we need a standards-based architecture that can offer a secure, flexible, and scalable clinical data sharing infrastructure for improving collaborative decision support. Blockchains have recently been touted as an innovative platform to guarantee transactions integrity and validity in financial and government sectors. Beyond its use in cryptocurrencies, blockchain has particular appeal for IoT health data given its emphasis on sharing, distribution, and encryption of patient's data across geographically distributed zones without compromising their content [32]. Blockchains enable trustless disintermediation of medical records making it possible for multiple healthcare stakeholders, who do not trust each other, to exchange digital assets while preserving their anonymity and their privacy from each other [21]. Thus, blockchain technology can allow multiple stakeholders to agree, at regular intervals, about the true state of shared data [12]. Such shared data can represent credentials and attributes of transactions, information about individuals, entities, etc. Health data could then be represented on the blockchain, and subsequently it could assign access rules, credentials, and permissions around those data to enable easier sharing.

Despite this promise, blockchain can be cost ineffective since it requires substantially high compute power and higher energy is required for creating trusted transactions [34]. Thus, blockchain scalability and decentralization are currently at odds as all IoT nodes need to store the entire blockchain transactions, state of account balances, contracts, and storage. As wearable computing devices may publish patient-generated health data (PGHD) to provide activity monitoring of personal health data, the number of connected IoT devices is forecast to grow to almost billions in the next decade. Specifically, the increasing use of wearable computing and telehealth programs in health IT infrastructure are consuming more and more of the network capacity to support more IoT devices, which in turn needs countless network equipment that should be maintained, managed, and upgraded individually to support the fluctuating network traffic. In other words, scalability becomes an issue particularly when it comes to processing billions of transactions that are expected on these devices [18].

Additionally, aggregating heterogeneous sensory data from different types of sources needs an agile infrastructure that embraces message brokers, sensor virtualization and softwarization for flexible, cost-effective, secure, and private IoT deployment for diverse applications and services [22]. To address these key challenges, there is a need to enhance the future blockchain-enabled healthcare industry through intelligence to enable a successful deployment and realization of powerful and trusted healthcare networks. Software-defined network (SDN) [20] and network function virtualization (NFV) [25] show a significant promise in meeting healthcare communication needs. SDN/NFV offers a flexible and programmable environment to support customized security policies and services in a dynamic and software-based fashion.

They also allow scaling up IoT capabilities by allowing on-demand service orchestration and management in healthcare domain [36]. In combination with blockchain, SDN/NFV can be used to develop and deploy privacy-preserving healthcare applications. They can optimize the flow management in response to attacks by enabling sophisticated traffic analysis and improving security and privacy based on global network awareness. SDN controllers can distribute security policies between the blockchain nodes and IoT healthcare network infrastructure. They can also enforce security and trust between IoT healthcare gateways and their local and body sensors as well as among distributed gateways.

In this chapter, we present a survey of the emerging trends and prospects on blockchains and SDN/NFV for health care, followed by an in-depth discussion of major key technical requirements for creating an interoperable health IT system. We then present BlockCare, which is our framework that fulfills the aforementioned requirements. BlockCare introduces a SDN-enabled blockchain architecture for interconnecting various IoT healthcare elements and blockchain nodes to control and manage patients' health-related parameters. We also present three use cases drawn from smart health care, precision medicine, and pharmaceutical supply chain, which will benefit from the proposed architecture. Finally, we discuss open challenges that must be addressed in the near future to support other technical requirements for improving advanced healthcare IT systems and highlight key lessons learned from our case studies.

4.2 Background on Enabling Technologies

This section provides a brief overview on blockchain and software-defined networking (SDN).

4.2.1 Overview of Blockchain and Its Use in Health Care

Blockchain is a decentralized distributed ledger that maintains a fully replicated database. When used for healthcare applications, it can store patient records and transactions at multiple nodes or sites without the need for centralized third-party authority [41]. The blockchain data layout is shown in Fig. 4.1, where transactions are represented by addresses, private and public keys, data, and hash values. Transaction data are used to store user records while hash values store coded and encrypted information generated from the previous block. The first block does not have a hash value to any previous block and is called the genesis block. The blocks contain data in a secure form, e.g., timestamps and hash tree (i.e., a Merkle tree) to store data blocks and hash the previous transactions.

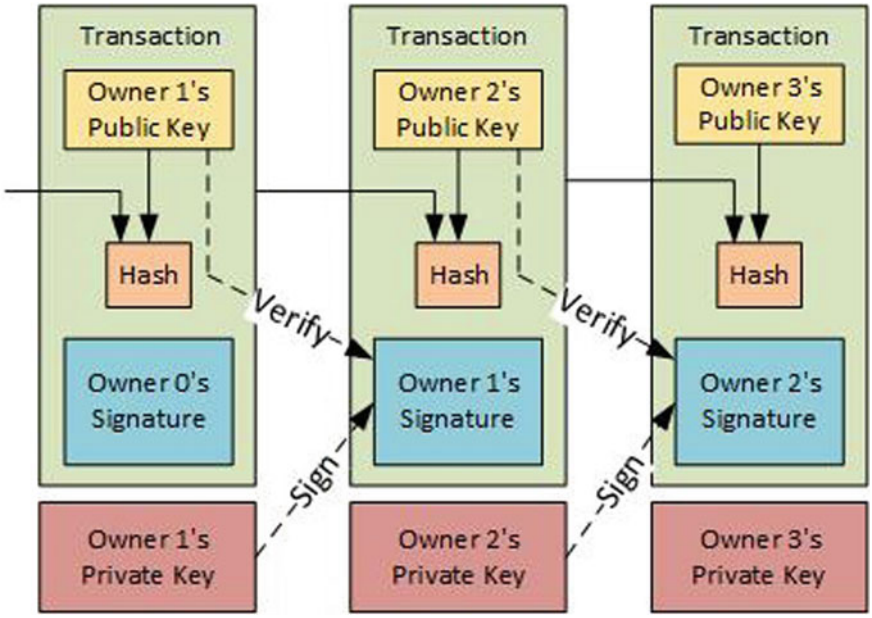


Fig. 4.1 Blockchain data layout

In the healthcare context, a transaction captures a specific healthcare service provided to healthcare entities, e.g., providers, payers, and patients. The transactions could include data, such as a patient ID, clinical data stored in electronic health records (EHR) systems, history and gaps in care from payers, and family history and device readings from patients [27]. Those data are stored in ordered blocks, which once recorded are resistant to modification and cannot be altered retroactively by design. All blocks are linked together to form a single sequence or chain that use cryptographic validation to link the blocks together. Each block references and identifies the previous block using a hash function which forms an unbroken chain. The chaining of the content means that any changes in the previous block require changes in all the latter blocks. Chain blocks can be replicated and distributed among all nodes in the blockchain networks, where the chained data are synchronized globally [37]. Each node in the chain stores patient records, historical, information, and control data. These data are executed locally at each participant node connected to the blockchain. This allows the system to achieve a global view of the patient's medical history in an efficient, verifiable, and persistent way.

Since the chained data must be synchronized, there should be a set of computer programs or protocols that embody agreement between two or more nodes [19]. Such protocols are called "*smart contracts*" and have the capacity to conduct intensive real-time tasks, execute heavy workflows, and can potentially be used for automation, reporting, and monitoring encrypted transactions. Smart contracts allow creating decentralized autonomous organizations (DAOs) and decentralized apps (DApps)

[38]. Smart contracts are autonomous and self-executing software code to control the exchange and distribution of digital assets between multiple parties according to certain agreements between the involved participants in the blockchain network while remaining is tamper resistant to any external control.

The blockchain is a distributed, transparent, immutable, validated, secure, and pseudo-anonymous ledger that exists at multiple nodes such that if *51%* of the nodes agree, then the trust of the chain is guaranteed. A distributed ledger itself can also be programmed to trigger transactions automatically using decentralized consensus to achieve large-scale and systematic cooperation in an entirely distributed and decentralized manner. The blockchain is distributed because a complete copy lives on as many nodes as there are nodes in the blockchain network. These nodes compete to validate the newest block entry before the other nodes to gain a reward for doing so. The information embedded in the healthcare transactions is encrypted and stored in the blockchain and can only be decrypted by authorized entities that have private keys. That is, the blockchain is immutable as none of the transactions can be changed, erased, or altered, i.e., all medical records, old, and new are self-auditing and will be preserved forever with no ability to delete.

The blockchain is validated by the so-called miners who are compensated for building the next secure block. Blockchain establishes trust to provide a way for all participants to have an up-to-date ledger that reflects the most recent transactions or changes. In the context of health care, GemHealth[1] is a platform that is built on blockchain to enable collaboration between different stakeholders into the sharing and transferring of healthcare data. The GemHealth blockchain supports pseudo-anonymity, where the identity of different healthcare operators involved in the transaction is represented transparently through a universal data infrastructure, i.e., address key. Blockchain can also be used for ensuring pseudo-anonymity of recruited participants for clinical trials [1].

Additionally, the private—so-called *permissioned blockchain*—can be used for creating a decentralized record management platform to allow patients accessing their medical records across multiple providers [3] while ensuring authentication and data sharing. A permissioned blockchain is also used for managing and sharing medical records for, say, cancer patient care [8, 9]. Blockchain transactions are created with a block of data comprising encrypted patient's identity (e.g., social security number, date of birth, names, and zip code) and aggregated meta-data including medical data files that should be stored inside remote cloud servers, whose access is managed by the blockchain logic. Zhang et al. [38] introduced a common software pattern to improve the design of blockchain-based health apps.

Similarly, blockchain healthcare platforms have been introduced [31] to ensure the trust management of medical data and foster patient-driven interoperability through data availability [12]. To satisfy interoperability needs, Peterson et al. [27] introduced a healthcare blockchain by encapsulating the Health Level Seven International (HL7) and Fast Healthcare Interoperability Resources (FHIR) standard (i.e., a draft standard API for describing the format of clinical data to exchange) for shared clinical data.

[1] https://enterprise.gem.co/health/.

To communicate healthcare context, they introduced a "*proof of interoperability*" based on Merkle tree blockchain structure as the consensus mechanism during block mining. The miners verify the clinical messages sent from the blockchain to ensure their conformance to the FHIR protocol.

Attempting to revolutionize the supply chain process of medicines, blockchain allows monitoring the movement of goods and drugs in real time [7]. It also simplifies the medicine inventory management and improves the efficiency of logistics by optimizing the transport resources, production, and distribution of medicines. Moreover, it helps in medicine warehousing, specifically to forecast the demand for medicines, historical sales data, and to build the life cycle of a medical product. Blockchains have also been used to reduce the operational costs in the pharmaceutical supply chain and assert health data immutability and public accessibility.

4.2.2 Overview of SDN and NFV

Software-defined networking (SDN) is a relatively new architecture to deliver dramatic improvements in network agility and flexibility [23]. The primary idea behind SDN is to move the control plane outside the switches and enable external control of data plane through a logical software entity called controller. The abstraction of the control plane provides common application program interfaces (API), i.e., southbound interfaces, to the underlying hardware to allow routing services, control, and management. OpenFlow is one of the most widely used southbound protocols. Additionally, the controller offers northbound interfaces (API) to enable efficient orchestration and automation of the network and align with the needs of different applications via SDN network automation and programmability. Along with SDN, NFV allows enforcing the security policies of the SDN network, by enabling the deployment of virtualized network functions inside virtual appliances [13].

4.2.3 SDN-Enabled Blockchain

There are recent efforts in integrating SDN with blockchain. For example, Kataoka et al. [17] combined SDN and blockchain to automate the process of doubting, verification, and trusting of IoT services to prevent them from attacks. The SDN controller receives information from trustable services and transform them into OpenFlow rules to program the underlying SDN switches. Similarly, Samaniego and Deters [30] virtualized IoT resources by combining blockchain and SDN to enforce permission-based communication during resource provisioning. Their approach improves network latency and throughput when blockchain is applied at the network edge. Likewise, Steichen et al. [33] proposed the ChainGuard framework as a SDN firewall atop the Floodlight controller to enforce blockchain security. ChainGuard filters the network

traffic and intercepts illegitimate packets to mitigate flooding attacks and prevent malicious behavior from vulnerable sources.

In the context of healthcare applications, Salahuddin et al. [29] proposed a softwarized infrastructure for supporting agile, flexible, and privacy-preserving deployment of IoT systems for healthcare applications and services. In this architecture, a software NFV manager is used to chain and orchestrate virtual network functions (VNFs) and help the SDN controller to steer the traffic among virtual and physical healthcare appliances. A blockchain holds a complete medical history for each patient with multiple granularity of control by patients, doctors, regulators, hospitals, and other healthcare stakeholders and provide a trust mechanism to secure patient records and keep all transaction blocks visible to a group of medical surgeons, doctors, and patients. The architecture guarantees the security of healthcare IoT gateways by protecting patient information and other tamper-aware medical records against any malicious traffic analysis and unauthorized data manipulation.

4.3 Key Technical Requirements for SDN-Enabled Blockchain

This section delves into the key technical requirements for creating a flexible, interoperable, secure, and tamper-proof healthcare IT system and analyzes the implications for SDN-enabled blockchain-based system design.

4.3.1 Requirement 1: Dealing with Fragmented Data

Figure 4.2 shows a typical data generated by a single patient visit where medical records are exchanged between different parties involved in the EHR system. In Fig. 4.2, data are created when a patient undergoes some tests, e.g., computerized axial tomography (CAT) scan, mammography, etc., and the results are to be disseminated to the radiologist and then a physician. Data could also be generated by wearable computing or monitoring devices to track the patient genomic or even data written by interviewing the patient. Subsequently, the EHR system should consider these actions to prepare an insurance claim. The result of the visit is stored in the hospital database and could be accessed remotely by a physician, while the claims and invoices are sent to the insurance. As a result, patient's data are fragmented across many institutions and data capture systems.

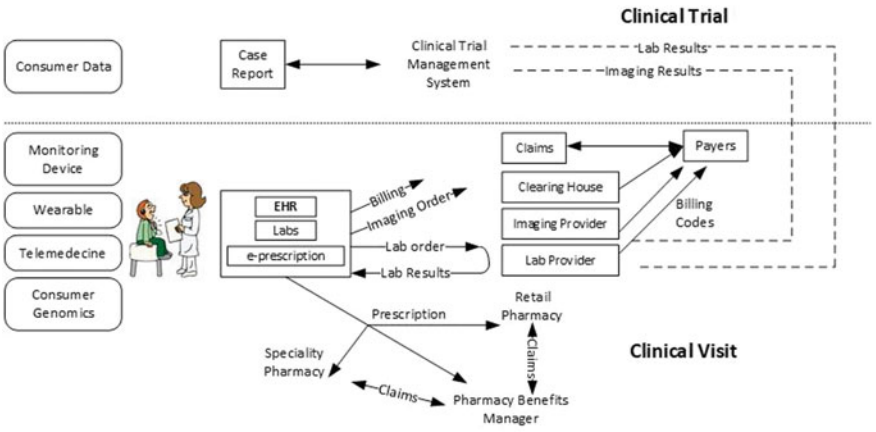


Fig. 4.2 Data exchange in electronic health records (EHR) system for a single patient visit

In such a scenario, practitioners, insurance companies, and providers end up using multiple third-party software tools to manage claims and thus must struggle with fragmented and siloed data as they are still using manual work to integrate different software. This is mainly because they continue to use delayed communication, vendor-specific care applications, and incompatible healthcare systems, which makes their coordination more difficult.

To provide patient-centered care, the healthcare industry needs to track the entire continuum of care and the billing cycle to reduce friction between the involved parties. Blockchain uses distributed ledgers to efficiently track transactions and claims and could help resolve claims instantaneously. Blockchain nodes can aid in providing accurate clinical data using decentralized storage, while data can still be shared across distributed nodes and networks. Additionally, blockchain permissioned ledger [35] with verified nodes can be used to restrict access to medical records to only authorized users who are tracked by public and private keys usage. Integrating all the healthcare industry in the same blockchain could allow coordinating permissions and access roles between various parties requesting medical records. A smart contract could include all the healthcare meta-data to directly control the exchange of digital medical records.

4.3.2 Requirement 2: System Interoperability

Exchanging medical records between the healthcare stakeholders requires significant collaboration between the involved entities to provide patient-centric care. Data created by business entities like hospitals, private clinics, and pharmacies are often soiled within the information system that creates it and results in creating scattered

individual patient's health data across numerous systems so that none of those institutions has a complete picture of the patient's data [39]. However, the exchanged medical information requires a consistent data format to be processed by complex patient matching computer algorithms and software. Despite some existing standards, such as HL7 and FHIR, that provide some basic interoperability for data exchange in healthcare information management systems, their level of interoperability is limited to some specific standards and the implementation of large-scale systems requires an extensive effort for mapping medical records to them. In particular, lack of interoperability spans to patient registration in medical systems, authorization procedure and government regulations, incompatible digital patient records, accounting, and billing, etc.

4.3.3 Requirement 3: Consistent and Secure Data Access, Storage, and Exchange

Recall that healthcare information often involves sensitive medical data that needs privacy protection to prevent any potential vulnerability that could reveal patient identities. Storing medical records in open peer-to-peer distributed blockchain nodes could be vulnerable even with the existence of encryption algorithms to hide data content because attackers could exploit the vulnerabilities of smart contracts and lock their behaviors. Thus, sensitive medical records should be stored and exchanged securely and kept tamper resistant to outside control.

Besides, medical data exist in different formats depending on the third-party applications it manipulates. Sharing these data among diverse providers, hospitals, etc., could not be meaningful to many of these healthcare stakeholders. Healthcare companies need consistent and rule-based methods to access and analyze data through smart contracts. Therefore, any blockchain decentralized apps (DApps) for healthcare system should consider existing standards for representing the clinical data. One approach to ensuring data consistency while keeping the healthcare secure and tamper-proof is to use modern data exchange formats such as JSON and protocols that support RESTful APIs.

4.3.4 Requirement 4: Identity Verification and Participant Authentication

Healthcare data contain sensitive and personal information that could be attractive to many parties. Therefore, healthcare systems require access control models to prevent cyber attackers from using sensitive data for financial transactions with third-party providers, who may perform data analysis to identify individuals [19]. Furthermore, the EHR system should ensure privacy and data integrity to protect the data from

unauthorized access attempts from inside the network so that the system grants access to only authorized parties. Conventional cryptographic primitives and access control models that empower the traditional IoT communication has reached their limitations and cannot ensure the protection of health data existing in the healthcare industry, specifically when attacks could be intentional.

4.3.5 Requirement 5: Timely Access, Cost Effectiveness, and Maintaining Modularity

Hospitals tend to be complex organizations, particularly when they also include research, teaching, and care services. Healthcare research networks often have specific requirements, which are different from the patient and the family access, who often need Internet connectivity. That is, information collected from heath data sources like mobile devices, patient's wearables, documents, EMRs, records, and images files should be shared on a timely basis to deliver appropriate treatment. On the other hand, the IoT devices and body sensors are expected to reduce the cost, increase the quality of life, and enrich the user's experience. That is, any static network configuration and monitoring become ill-suited to support the future healthcare applications. Therefore, healthcare providers need flexible and modular network infrastructure to support the growing healthcare traffic. Such an infrastructure should offer a cost-effective solution for improving agility and flexibility of sensor networks. Similarly, a unified voice and video communication is crucial for virtual medical interactions and exchange of records and imaging files so that the communication could be performed at the proper levels of quality of service (QoS).

To fulfill these requirements, blockchain allows updating the transactions about the patient's data in near real time, thanks to the distributed ledger system which facilitates the distribution of updated records, while ensuring secure access to health data. Blockchain smart contracts should be modular by design to avoid information lock-in due to the immutability of smart contracts. Smart contracts should be loosely coupled with other components in the healthcare system so that every modification on it should not create a new instance of the smart contract to avoid frequent upgrades among involved participants. Thus, blockchain helps in avoiding the use of third-party applications, which offer no time lag for accessing data. Besides, combining blockchain and SDN could improve the network scalability by increasing the number of involved blockchain nodes and adapt the network condition to the frequent changes in user's locations. Integrating both technologies could make it possible to manage the entire network through intelligent orchestration and provisioning systems.

4.4 BlockCare: SDN-Enabled Blockchain for Scalable and Tamper-Resistant Healthcare Network

This section delves into the architectural details of our solution called BlockCare and shows how it addresses the aforementioned requirements. In particular, we show how our approach enables supporting scalable, dynamic, and flexible resource management with our SDN-based framework and present the algorithms to perform tamper-resistant IoT on blockchain communication in symbiosis with SDN.

4.4.1 BlockCare System Design

Figure 4.3 illustrates BlockCare, which is a softwarized agile, flexible, cost-effective, tamper-proof, and privacy-preserving framework for IoT-enabled healthcare applications. Healthcare DApps are SDN-aware as they can listen to mining nodes and report suspicious IP addresses and validate known packets. The BlockCare framework introduces a *proof of ownership (PoO)* concept for IP addresses and reports them under smart contracts. It also supports an intrusion detection system in the

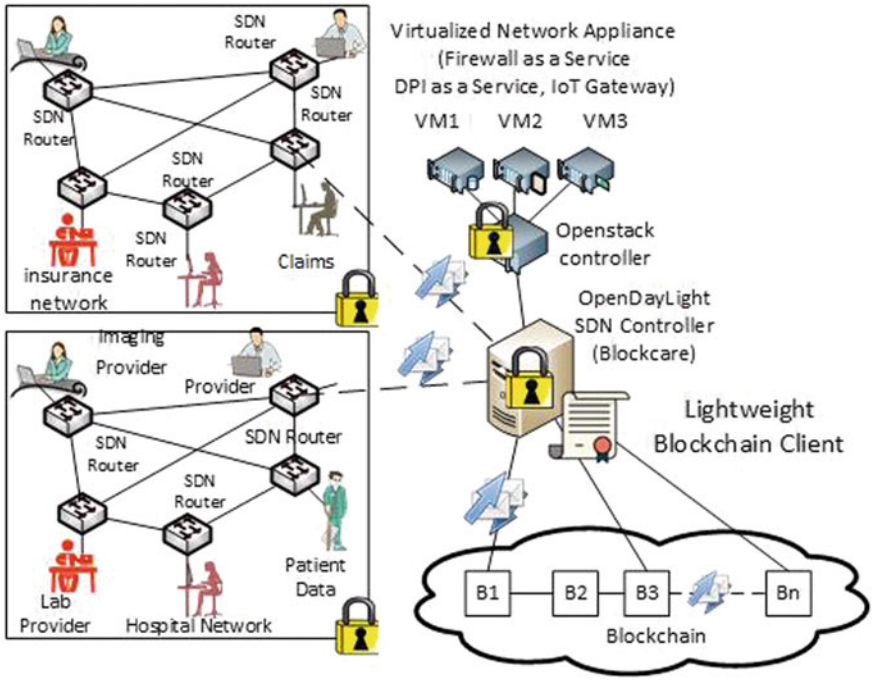


Fig. 4.3 BlockCare system design

form of virtualized network functions (VNFs), i.e., a Firewall as a Service in SDN OpenFlow-based network, which takes care of malicious flows and enables DDoS detection and mitigation on demand.

BlockCare delegates storage of blacklisted and whitelisted IP addresses in virtualized network functions (VNFs) instances inside trusted Docker containers to maintain all the reports about whitelisted and blacklisted IP addresses. VNF instances can be dynamically deployed to meet changing conditions and accommodate higher traffic demand or more stringent service requirements. Furthermore, our approach introduces a proof of ownership (PoO) of IP addresses and reports them under IoT smart contracts. It implements a blockchain DApps to enforce trust on transactions by listening to mining nodes and report suspicious IP addresses to the SDN controller which should validate known packets (i.e., unknown packets are considered suspected and are dropped).

Figure 4.4 illustrates how the SDN controller can enforce IoT data exchange with the blockchain. To offer a better level of adaptability and dynamically respond to possible threats/attacks, the SDN controller contains a smart verifier to perform network topology management and provide access control mechanisms, authentication, traceability, and permission services. Additionally, the SDN controller implements security policies in the control plane to protect the SDN router against eventual intrusion. As the SDN routers are directly connected to the blockchain, the data should be encrypted before its transmission to remote participants.

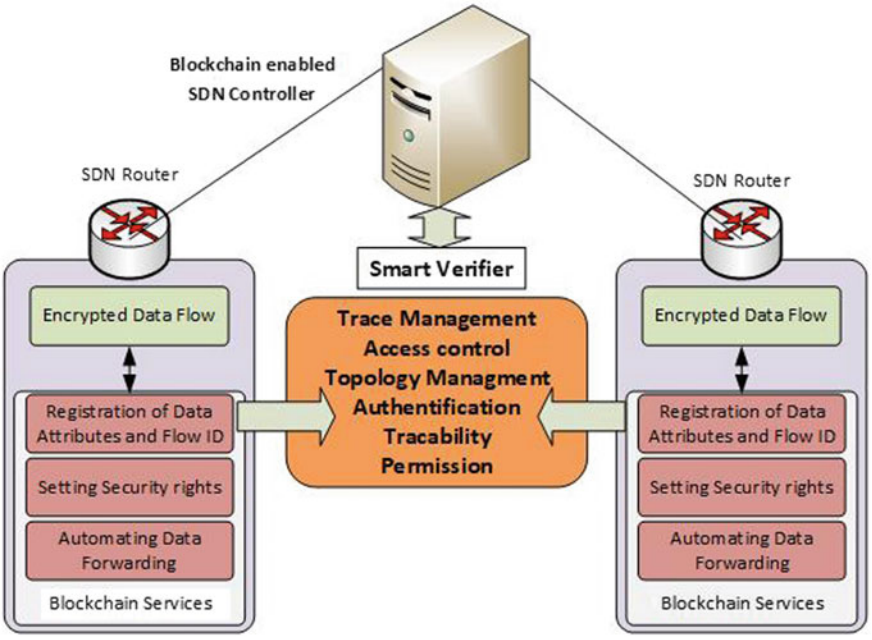


Fig. 4.4 Securing the data plane transactions

The steps are as follows: First, the SDN controller registers meta-data and their attributes, and creates flow ID for those flows. Then, it classifies packets into several classes to create the required security rights for each class of packets. Finally, it coordinates with the SDN router to perform the same behavior and automates the processing of all the arriving packets belonging to those classes. Once the medical records are encrypted, transactions can be sent to the blockchain in a secure fashion.

4.4.1.1 Addressing Requirement 1: Data Aggregation to Address Fragmented Data

In the current electronic health record (EHR) systems, a patient can access a given healthcare institution using their own portals. The patient uses his/her credentials to authenticate on the system and perform the operations (i.e., update, modify, delete his personal details, etc.). BlockCare allows creating a national patient's identity so that it associates different clinical records of a single patient located in multiple healthcare entities into one individual account. This is particularly important for several countries such as USA where there is no national patient identification. Individuals can use their own public/private keys to identify themselves in the healthcare blockchain. Thus, medical stakeholders can use the client public key to identify his national ID and share medical information with other providers, patients, and their proxies. By adding patient's public keys on the system, individuals can also add new permissions to securely grant customized/personalized access to their medical data or the meta-data on the blockchain. Additionally, as a patient could have different IDs across multiple institutions in existing EHR systems, BlockCare can resolve this issue by enabling data aggregation of patient medical records using their unique public keys. BlockCare uses an individual public key for each patient so that it could be used to resolve any conflicting identities between patients.

4.4.1.2 Addressing Requirement 2: Enhancing Patient-Centric Interoperability

Interoperability enables providers to securely and scalably share patient medical records with one another regardless of provider location and trust relationships between them. As interoperability becomes patient-driven, BlockCare provides patients greater control of their own data and can extend it to their proxies through data availability. It introduces the concept of *pegged-sidechains* [4] (i.e., private blockchain) for coordinating the data exchange between involved parties and streamlining the flow of information for medical records. Domain-specific sidechains could rapidly grant authorized access and privileges to the involved healthcare stakeholders and synchronize with the main blockchain.

Furthermore, BlockCare could improve operational effectiveness by reducing administrative tasks, e.g., manually entering medical data to the system after receiving paper. It also helps in avoiding duplicated clinical interventions, decrease

waste management, and reduce exposition to radiation procedures. Second, the use of meta-data to communication healthcare information in a form of structured data facilitates the access to relevant clinical information. Finally, it also offers a collaborative clinical decision support as structured data are easy to manipulate and interpreted by the involved healthcare stakeholders.

4.4.1.3 Addressing Requirement 3: Improving Security and Consistent Data Exchange

In a blockchain principal, the data included in the smart contracts are often open to public and available on the Internet so that users can simply access them by selecting the right URL. Data representation in healthcare application is more complex and more sensitive to storage in the public blockchain. In particular, security and privacy are critical issues since we will perhaps see new algorithms to decrypt data circulating across blockchain networks. To avoid such a situation and further enforce the security of medical records, BlockCare keeps sensitive medical data of chain rather than storing them in open healthcare blockchain system. In this case, we can keep references to meta-data referencing protected data which helps the providers in maintaining the ownership of their medical records. Moreover, NFV creates virtualized security functions inside virtual appliances to enforce the security at the network edge. Thus, it could prevent attackers from intercepting patient-sensitive data stored at the provider's network.

Furthermore, to improve data consistency on the healthcare application, BlockCare uses the JavaScript object notation (JSON) data format, which is supported by most web services and can easily be consumed by all programming language. The JSON format can easily be used by different blockchain clients which use JSON-RPC for their data transmission. Therefore, all BlockCare transactions can be encoded in the JSON encoding to make them readable at the client machine.

4.4.1.4 Addressing Requirement 4: Security Enforcement for All Participants

Traditional SDN security approaches deploy whitelists, i.e., SDN firewalling, to enforce trust between the SDN controller and the underlying devices. The authentication is based on identity to ensure impersonation prevention, protect the control and data planes against intrusion, and ensure that malicious attacks do not tamper with the controller configuration. As shown in Fig. 4.5, BlockCare enforces the security and the authentication for each involved healthcare entities by offering integrity on those firewalls. It also employs public/private key cryptography to create digital signatures and encrypt data. The public key identifies the user's identity, while the private key allows the user to access the blockchain network from their remote locations. Additionally, BlockCare offers new cryptographic mechanisms to ensure the robustness of both the control plane and data plane. It employs the blockchain Merkle tree [11]

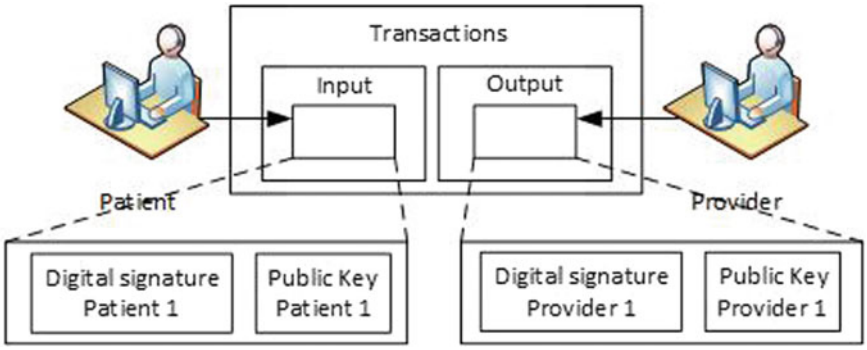


Fig. 4.5 Exchanging secure transactions using the public key

to regenerate the root hash from any given leaf in the tree, i.e., it could provide the proof of participation of all SDN nodes (i.e., any involved healthcare entity) in the network and ensure that malicious nodes are not be able to impersonate real SDN ones.

Furthermore, BlockCare supports the concept of M-of-N signatures (i.e., Multisig digital signatures), where multiple associated private keys could be used to decrypt the data. This is particularly interesting for virtual medicine interaction, where members of the medical team often use at least M of their private keys to sign, encrypt, or decrypt healthcare transactions. This is also useful for incapacitated patients to grant some of their private keys to authorized caregivers and doctors to grant access if they cannot provide consent to access the data.

4.4.1.5 Addressing Requirement 5: Softwarization and Virtualization

Figure 4.6 illustrates how BlockCare uses SDN/NFV to create a set of network functions that can be deployed into software packages, assembled and chained to create the same services provided by legacy networks. By leveraging BlockCare, the SDN controller will be able to sign and verify health transactions across distributed nodes in which data could be signed and verified in near real time. The softwarization and the virtualization of the provider's network infrastructure allow provisioning virtual health IT appliances on a generic hardware. Adopting these principles in healthcare communication can help reduce capital expenditure (CAPEX) and operational expenditure (OPEX) and enhance the performance and the QoS delivered to the provider's networks.

Furthermore, since individual IoT devices, wearable computing, and body sensors are not so powerful to perform blockchain mining, SDN/NFV can enforce their coordination and performance by creating a modular architecture in which virtual miners can be hosted inside a NFV platform such as the open platform for NFV (OPNFV). The BlockCare framework can dynamically provision virtual appliances,

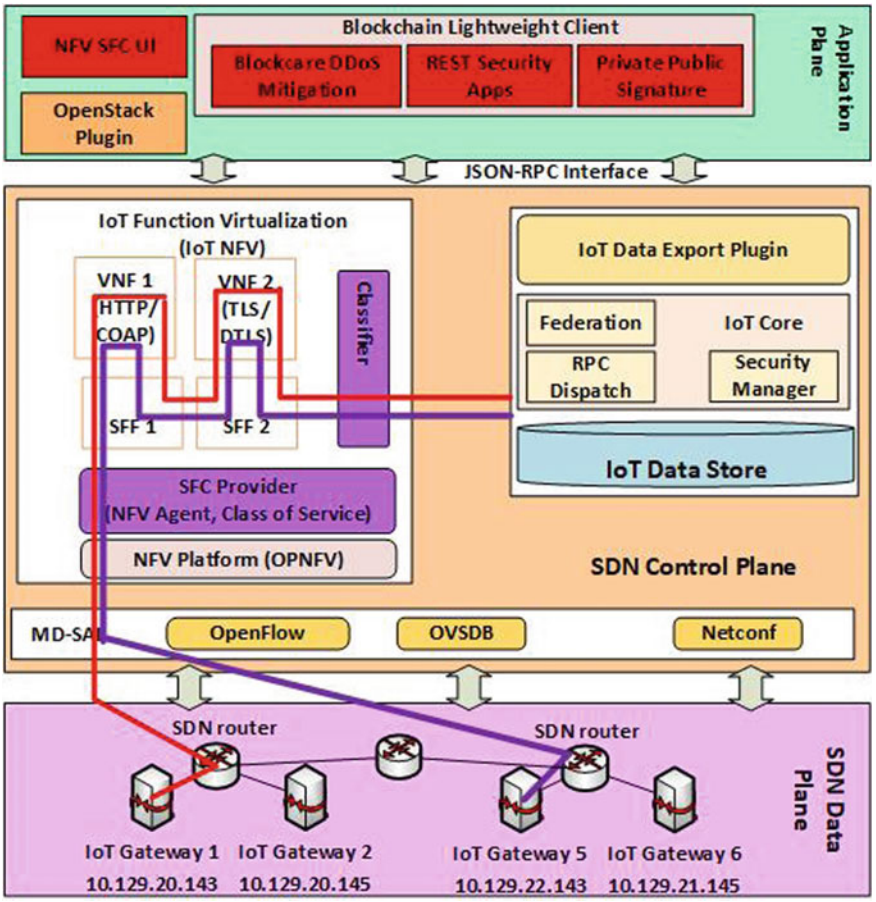


Fig. 4.6 BlockCare's proposed SDN/NFV solution on the blockchain architecture

scale them up/down according to the needed mining performance. On the other hand, BlockCare protects SDN security policies and configuration and flow table rules from intentional or unintentional tampering. As a result, the SDN northbound and southbound interfaces (API) can be more accessible, and the network control information can be consolidated into a few number of locations instead of being spread over the entire network.

4.4.2 Smart Contract

BlockCare allows trusting the IoT devices used in the healthcare environment based on their MAC and IP addresses and their interaction with other IoT services. As

shown in Listing 4.1, a data structure (Lines 29–38: SuspectBehavior) inside the blockchain smart contract is used to detect suspected behavior and reports (Lines 23–26: Report) it to the SDN controller. This latter can now distribute trusted lists of IoT devices, body sensors, wearable computing, etc., connected to patient home environment and share their data with the other healthcare stakeholders across the Internet. Therefore, a blockchain validator is introduced to check the validity of IoT devices connected over the blockchain. The validator parses the OpenFlow messages

```
1
2  pragma solidity ^0.4.24;
3
4  contract Blockcare
5  {
6  address public owner;
7  IPAddress ipBoundary;
8  Lifeliness currentState;
9  Lifeliness constant defaultState = Lifeliness.sensing;
10
11 enum Lifeliness {awake, sleeping, poweredOff, sensing};
12 event CommunicationFailure();
13 event TriggerAlarm(string src_IP, string src_MAC, string data);
14 event DiscoverElement(string IP, string MAC);
15 event Sent(address from, address to, uint amount);  // event to
       allow client to react on changes efficiently
16
17
18 mapping (address => Misbehavior[]) public MisbehaviorList;
19 struct IPAddress {
20 string ip;
21 string mask;
22 }
23
24 struct Report {
25 uint expirationdate;
26 IPAddress sourceIp;
27 IPAddress destinationIp;
28 }
29
30 struct SuspectBehavior {
31 address subject; //subject who performed the misbehavior;
32 address object; //
33 string res; //
34 string action; //action (e.g., "read", "write", "execute") of the
       misbehavior
35 string misbehavior; //misbehavior
36 uint time; //time of the Misbehavior occured
37 uint penalty; //penalty (number of minitues blocked);
38 string suspectIP;
39 string suspectMAC;
40 }
41 }
```

Listing 4.1 Smart contract for detecting malicious IoT devices

to identify the source and destination of incoming traffic. The SDN controller uses this information contained in the OpenFlow packet headers to create a wide network view including topology state and transactions meta-data.

By expecting and parsing every OpenFlow packet exchanged between the IoT devices and the network, the SDN controller can identify every abnormal behavior in the network. That is, if an attacker wants to take control of any IoT device, the changes of the device ownership in the network will be visible in the topology viewer module within the SDN controller. This method allows the SDN control plane to distinguish two types of lists, i.e., blacklist devices and whitelisted ones. The former are the suspicious users whose behavior is abnormal (i.e., representatives of malicious attack or unexpected behavior) so the controller should isolate them from sending traffic on the blockchain. The latter are the users or devices whose behavior is normal, and they could continue delivering their content as they belong to the blockchain.

To further enforce the security of the exchanged transactions, Listing 4.2 illustrates how we can restrict a set of IP addresses to a limited number of involved IoT participating devices on the blockchain. If the SDN controller detects any malicious traffic coming from suspicious IoT devices, it can remove it, or even remove the IoT gateway from the blockchain network.

```
2  function ArrayStore(string ip, string mask) needsMask(mask) {
3  owner = msg.sender; // address of the creator of the contract
4  ipBoundary = IPAddress({ip: ip,mask: mask});
5  }
6
7  function createCustomer(address customer, string ip, string mask)
       public ownerOnly (mask)
8  {customers[customer] = IPAddress(ip, mask);}
9
10 modifier isInSameSubnet(string ip, string mask) public constant
       returns (bool success)
11 { if(mask < ipBoundary.mask) returns false; }
12
13 modifier needsMask(uint8 mask)
14 { assert(mask == 0);    _; }
15
16 modifier ownerOnly()
17 { require(msg.sender == owner); _; }
18
19 function removeM2MDevice(string MAC) public ownerOnly returns (
       bool success)
20 {   bytes32 key = keccak256(MAC); //Use encryption
21   delete M2M_Device[key] ;}
22 function deleteM2M_Gateway() ownerOnly returns (bool success)
23 {   bytes32 key = keccak256(MAC);
24   delete M2M_Gateway[key];}
25
26 function deleteAll_M2M_Gateway() ownerOnly returns (bool success)
27 {   for (){ delete M2M_Gateway [key];}}
```

Listing 4.2 Smart contract for detecting malicious IoT devices

4.5 Use Cases

This section describes three case studies that utilize our proposed BlockCare solution in three different applications involving smart health care, precision medicine, and pharmaceutical supply chain.

4.5.1 Blockchain Technology for Smart Healthcare Clinical Trials

We assume the general case of digitized clinical trials where patients use body sensors and wearable computing from their homes to gather relevant health diagnosis data and send them via Internet to clinical research centers as illustrated in Fig. 4.7. We focus on the recruiting phase of participants for the clinical trials [40]. The approach enables bringing specific patient's profiles into the recruitment process based on predefined inclusion–exclusion criteria and biometric attributes collected from patient body sensors. Individual patients expect that the privacy and the confidentiality of their personal health data will be guaranteed during clinical trials.

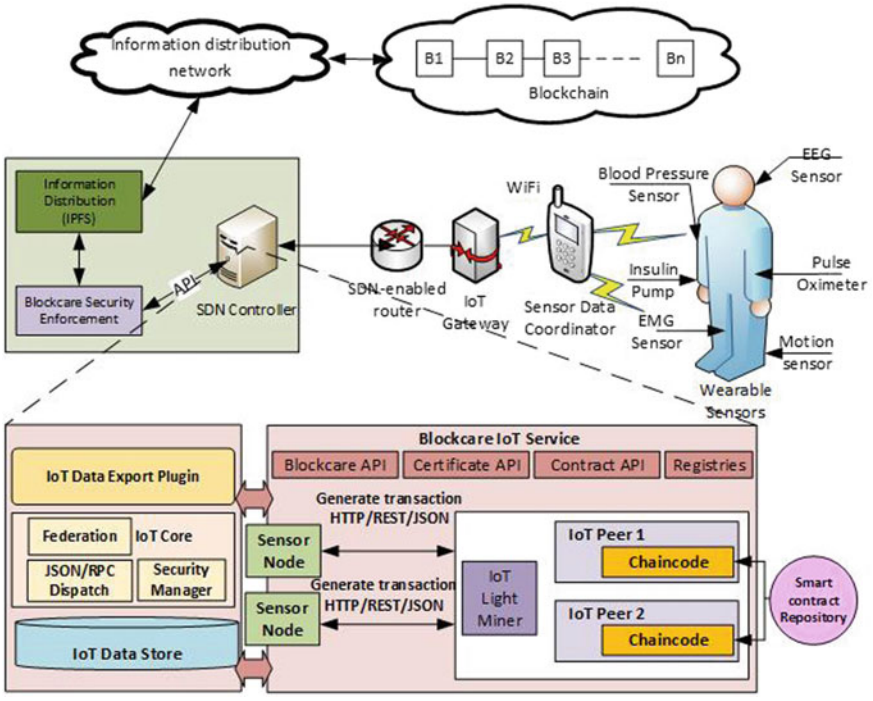


Fig. 4.7 Building smart health clinical trials

Figure 4.7 shows how BlockCare can be used in smart healthcare clinical trials. The wearable devices deliver information about the current health of the patient along with other data related to the manufacturer details (i.e., devices' model number, manufacturer ID, etc.). The handheld device is equipped with decentralized application (DApps) on the blockchain where the patient provides his/her public key to be recognized by the healthcare providers. BlockCare holds the SDN controller which contains the blockchain IoT service (including the blockchain API, certification API, contract API, and registries) and a lightweight miner to verify that transactions within each block are legitimate. BlockCare offers an IoT data export plug-in for simplifying the transmission of IoT healthcare information to the blockchain network using the JSON/RPC data exchange format. This method increases the security, efficiency, and consistency of the communication while keeping the design of edge network (i.e., the IoT gateway in the patient home network) simple. This approach of modularity in BlockCare allows research institutes to identify accurate data available from remote patients' devices due primarily to IoT gateways which can verify the integrity and authentication of every connected device and discard data coming from untrusted devices. Thus, the research institutes can easily analyze historical datasets provided by individuals and decide if the candidate could be included or excluded from the tests.

4.5.2 *Blockchain Technology for Improving Precision Medicine*

Precision medicine, also known as personalized, molecular, predictive, and preventive medicine, is a new scientific way to treat and prevent illnesses tailored to an individual based on a person's genes, family and medical history, and environment [31]. As illustrated in Fig. 4.8, it intends to deliver targeted treatments for each patient based on its unique genetic background. To drive further breakthrough treatments, researchers should be able to access the enormous genetic databases to identify

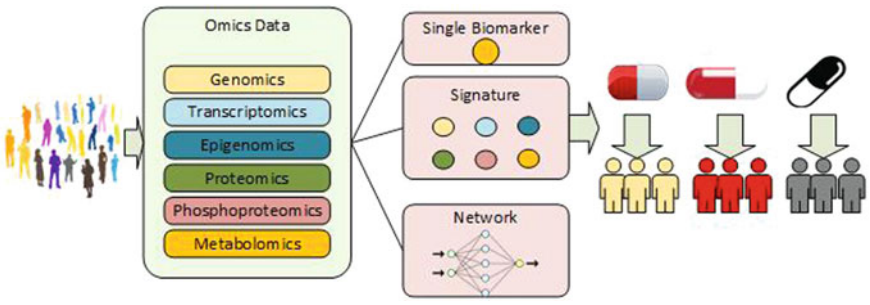


Fig. 4.8 Building precision medicine solutions using BlockCare

biomarkers based on DNA and genome sequences of millions of patients who have undergone some manner of genetic tests.

However, most genome sequences generated for clinical researches are stored in private and secured databases held by universities and research centers, which makes data sharing with other institutions impossible. The BlockCare framework could facilitate the implementation of precision medicine by allowing openness of genetic database to other researchers while ensuring the integrity, privacy, and security of their content. BlockCare keeps accurate medical records of all the previously generated transactions and allows preserving original ones intact. It can generate new transactions including DNA and genomic information from the original ones. The newly created transactions can be used to predict actionable, genetic mutations and assist in the drug discovery process. Thus, BlockCare could help precision medicine in designing personalized pharmaceuticals to target a specific patient's disease and giving individuals ownership of their genomic information, while preserving the medical datasets security and integrity against any malicious intrusion.

4.5.3 Blockchain Technology in Pharmaceutical Supply Chain

The logistic complexity of pharmaceutical supply chain operations is exacerbated with the fragmented and siloed data systems used by different healthcare stakeholders, i.e., pharmaceutical manufacturers, wholesalers, retail pharmacies, health facilities, etc. On the other hand, the global counterfeit medicine market, i.e., lifestyle, cosmetic products, and life-saving drugs, is exploding in the past few years (i.e., about 30% of the global drugs market) [24]. This creates opportunistic threats leading to economic and human costs, which needs effective collaboration and harmonization between trading partners and regulatory agencies to secure and optimize the pharmaceutical supply chain.

Figure 4.9 shows the current healthcare supply chain model (i.e., Fig. 4.9a) and the BlockCare's expected new model (i.e., Fig. 4.9b). The small circles in Fig. 4.9b describe a blockchain component that mediates data dissemination between the healthcare stakeholders. Those components can be considered as connectors between different medical professionals' databases involved in the supply chain because they can ensure the interoperability between clinical data silos. That is, BlockCare allows all trading partners better control over their inventories, automate their processes, record all their transactions, and save costs.

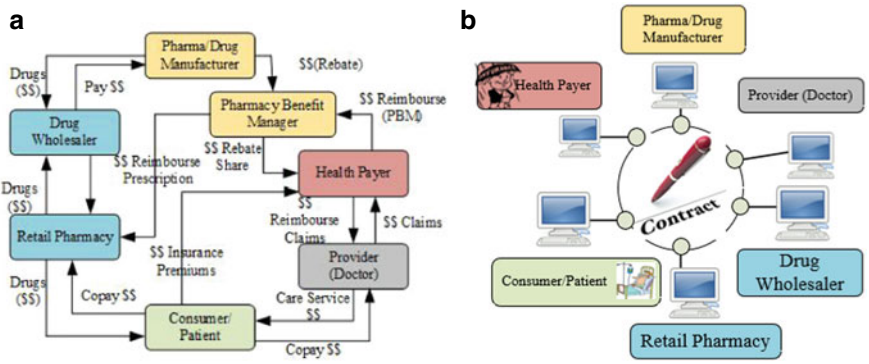


Fig. 4.9 Healthcare supply chain: current model versus BlockCare model: **a** current healthcare supply chain model; **b** BlockCare-enabled healthcare supply chain model

4.6 Open Challenges and Ensuing Directions for Future Work

Blockchains are heralded as the technological breakthrough to address serious security and privacy concerns and allow new applications and services to emerge. However, the progress of this nascent technology is accompanied by a complex set of unique challenging issues that require a major redesigning work for the network to develop infrastructure, agreements, and safety mechanisms to overcome them. This section delves into the key technical and government challenges that developers, regulators, and healthcare providers face and discusses some initial solutions to address them. These issues then become the key directions for conducting future research.

4.6.1 Scalability Issues

Blockchains are difficult to scale by design, and this is why cryptocurrency is not replacing fiat currencies and cannot be used as payment method like VISA or Mastercard. Indeed, Bitcoin can process four to seven transactions per second (tps) and Ethereum around 15, while the VISA credit card can routinely handle 2000 transactions per second (tps) and can accommodate peak volumes of 10,000 tps. This way we can use cryptocurrency to buy medication from a pharmacy and wait for an hour for the transaction to settle. The main reason for this slow speed is that every transaction and every computation should be performed at every blockchain node in the network, which limits the processing capabilities of the entire network to the processing capability of a single node. Smartphones have limited processing, memory, and computation capabilities and currently cannot be used as blockchain nodes to perform electronic payments like they do with NFC. Blockchain nodes

require powerful computers with massive memory, CPU processing, and bandwidth resources to process a few transactions.

Additionally, blockchain transactions undergo large latencies to process, e.g., the block time (the period for collecting transactions) in Bitcoin is 10 min and 15 s for Ethereum, respectively, which means that a transaction takes at least 10 min to be confirmed. On the other hand, for large transfer of amounts, it is recommended to wait more time (about an hour) to confirm the transaction and verify its integrity because it must outweigh the cost of a double spend attack, while the same processing takes seconds at most with VISA.

4.6.2 Power Consumption

One of the expensive costs in blockchain is the cost of electricity, which is needed for running both the hardware and the cooling systems for the miners. Mining uses massive amounts of energy, and all of it is wasted. Some recent estimates showed that the distributed consensus based on the proof of work consumes around 61.4 TWh, which is equivalent to 1.5% of electricity consumed in the USA. That is, validating and sharing one transaction require the same amount of electricity for powering 1.57 American households for one day ($600 million in 2017) and tend to consume as much electricity as Denmark by 2020, and all these costs are paid with fiat money, which imposes more pressure on the digital currency value. Additionally, this cost for making transactions trustable is very wasteful for mining. Mining involves competition against other miners, especially if we know that all blockchain nodes spend their resources with no benefit other than mining, as they compete for useless proof of work effort in hopes of the possibility of reward.

Besides, as the mining hardware generates a lot of heat, it may overheat the overall machines, reduce their efficiency, and even damage them. Conversely, off-the-shelf miners do not invest largely in cooling systems and air conditioning. Thus, it becomes more interesting to invest in submersion cooling where mining chips could be submerged in a cooling liquid that circulates to dissipate heat. Immersion cooling could help in packing more chips in the same mining hardware, i.e., packing more processing power and memory in the same physical space to increase the mining potential. Some other approaches tend to use under the sea cooling to develop self-sufficient underwater mining data centers.

4.6.3 Storage

The blockchain distributed ledger usually stores transactions and blocks in text or meta-data files. The size of each block depends on the synchronization mode activated at each involved participant node. For example, the blockchain size is closer to 50 GB for Ethereum using the “fast sync” mode, and this amount tends simply to grow by

14 GB each year. These heavy-sized data are the consequence of ensuring the data availability by keeping the same database with all nodes. As a result, it takes a very long time for a node to download the entire database (e.g., 1 day). Again, as the comparison with VISA, to handle the same amount of 2000 tps it would be 1.42 PB per year or 3.9 GB per day, and at 150,000 tps, the blockchain would grow by 214 PB per year.

Although some researchers believe that reducing the blockchain size to a smaller block could improve storage, it however could impact their scalability as bigger blocks can be more efficiently accessed. We need to rethink innovative compression algorithms to make blockchain storable at large scale, perhaps by inspiring approaches from modern big data and data-intensive science era. We also need to rethink improving the data accessibility using specific APIs that facilitate automated calls to the blockchain, obtain address balances and balance changes, and notify user applications when new transactions or blocks are created on the network.

4.6.4 Privacy Leakage

The key security issue in blockchain is the possibility of a 51% attack in which one mining entity could grab control of the overall blockchain and double spend previously transacted coins into his own account. This issue comes from the fact of centralizing the mining power into only few large mining pools who control the majority of the transaction recording. Double spending is also more vulnerable as it leads to spoofing users to resend transactions thereby allowing malicious attacks to double-spend coins. For example, in 2015, attackers have injected malicious traffic to the Bitcoin network and cracked existing transactions.

Another important security issue concerns the vulnerability of smart contracts themselves. The DAO attack is a dramatic example of what could happen when a smart contract is hacked. Hackers were able to move $60 million worth to an intermediate holding contact where they were in sole control. The attackers unleashed a malicious intermediate contract and invoked it repeatedly to perform memory-intensive work, thereby they dramatically slowed down the overall blockchain network. This happened because smart contracts have poor design and suffered from the reinjection vulnerability. They were designed without rolling back in mind so that the stolen amount was not recoverable. A possible alternative to overcome this limitation would be to make the design of smart contracts more difficult for bugs to appear. This will need formal verifiable language and model-checking tools for exploring all possible states and transitions in the smart contract models, which in turn requires domain-specific abstraction techniques to perform formal verification.

4.6.5 Smart Contract Risks and Governing Regulatory Status

Smart contracts raise a number of challenging legal issues. First, blockchain smart contract is autonomous software code that can run across distributed nodes that agree to its content. This creates some confusion with manual contracts, e.g., word contract with terms and conditions, as a smart contract depends on a number of factors and use cases that may be a barrier in establishing a legally binding contract compliant with the applicable law, i.e., cannot give rise to legally binding contractual relations under the laws. For example, the legal requirement of "SHOULD" or "certainty" in manual contracts may not be easily satisfied as not all smart contracts could operate in conjunction with the natural language contract terms.

That is, when a smart contract has a legally binding contractual effect, the technology within which it is deployed may sometimes give rise to problems in relation to legal enforceability. There may be no central administering authority to resolve a dispute or to recognize liability for mistake, error, or fraud. Dispute resolution mechanisms could address enforceability and jurisdictional variations. Inserting a dispute resolution mechanism into a smart contract will be pro forma to address the issues around enforceability and jurisdictional variations.

That said, several questions arise, e.g., when should the involved parties to a smart contract receive a notice of its terms and conditions and agree it, prior to or after its execution? How are these agreements required by law to be written, e.g., using computer code? Or, what are the changes that should be introduced to unwind transactions that should not have happened? If something goes wrong with the execution of the smart contract or someone suffer loss, where do they go for recourse? Certainly, the answer will vary significantly depending on the jurisdiction.

There are in fact two models of smart contract agreements being discussed. One model promotes the "code is contract" approach, while the latter considers it as digitizing the performance of business logic, e.g., payment, which may or may not be associated with a natural language contract. In between there is a novel jurisdictional variation emerging to connect the natural language terms of smart contract to computer code via parameters. Nonetheless, contracts sometimes deal with specific conditions that cannot be translated easily into "if this, then that" computer code. Force majeure is a good example, where manual contract often includes concepts of subjective judgment, reasonableness, and acting in good faith. These concepts cannot be translated into software code, neither connected to computer node via parameters.

4.7 Possible Future Directions

4.7.1 Off-Chain Computation

Off-chain computation is a promising technique to scale up the blockchain by realizing heavy tasks outside of it and only store the outcomes on blockchain. A key

to success of this approach is that clients can not only retrieve results and proof of correctness for an outsourced operation, but also verify the proof of correctness themselves with less computation power. The heavy computation is delegated to another layer on top of blockchain that performs heavy, compute-intensive work and wisely uses the blockchain. The off-chain computation could be centralized, distributed across a group of nodes, or outsourced in a sidechain. The original blockchain is referred to as the "*main chain*," and all additional independent blockchains are referred to as "*sidechains*". It helps to scale up the main chain and limits and isolates any damages to sidechain to prevent the main chain from any dramatic damage. Off-chain states are often private and confined among a group of participants, so that all transactions are cryptographically secured by using MultiSig scheme or a smart contract-based locking to enforce security.

4.7.2 Big Data Analytics

Blockchain can be combined with big data to transform the health care [6]. Currently, big data is stored in siloes within different doctor's computer systems, hospitals, and clinics. Integrating blockchain into big data analytics could transform the health care by predicting epidemics, curing diseases, improving the quality of life, and avoid preventable deaths [6]. Big data analytics can also reshape the data structure in healthcare services as it includes a huge amount of data gathered from body sensors and provide tools for performing relevant health diagnosis and monitoring and thereby help in making more prediction with patient's data and making recommendation to patients [16]. For example, BigchainDB[2] builds on top of big data distributed database to avoid any hard limit on the transaction size. This is extremely interesting for EHR systems to utilize the transaction meta-data to disseminate trained machine learning models and meta-data information model to involved healthcare stakeholders. Similarly, MedRec [3] blockchain builds on a data mining scheme to sustain medical records to medical researchers. Additionally, using off-chain computation we could have data management sidechains for storing patient's data as it is distributed among the medical community and involved healthcare stakeholders and a main chain to perform data analytics.

4.7.3 Artificial Intelligence

In light of the recent advances in blockchain technology, artificial intelligence (AI) could help in solving some challenging issues in blockchain, such as energy consumption, scalability, security, privacy, efficiency, and mining. DeepMind AI has been proven to be very efficient in optimizing energy consumption as it consistently

[2]https://www.bigchaindb.com/features.

achieved a 40% reduction in the amount of energy used for cooling in Google data center, so we believe that similar results could be achieved with blockchain as well. AI could also help in producing decentralized intelligences (either on-chain for basic information or off-chain in case of extra attachments) by introducing decentralized learning system such as federated learning and supporting sharding techniques to make the system more efficient.

4.7.4 Smart Contracts

Smart contracts are software code where bugs could exist and may be vulnerable to malicious activity. The DAO contract suffers from the reinjection vulnerability which was exploited by hackers to withdraw existing funds repeatedly. Therefore, smart contract performance analysis is very important as it will reveal the limitation and show the potential vulnerabilities that may occur. Thus, formal verification of smart contracts is an important task to detect any irregularities in its design and behavior. Model-checking approaches [26] and formally verifiable language and formal verification tools are necessary to verify that the smart contract implementation complies with its specification, verify its security properties, and formalize it by a set of temporal logic propositions [5]. At the time of writing, the Enterprise Ethereum Alliance and Hyperledger project formally announced[3] that they are working to identity an open blockchain architecture along with its open-source reference implementation compliant with Enterprise Ethereum Alliance (EEA) specification [10] to further improve the performance and the scalability of blockchain ledgers as well as providing more secure and interoperable smart contracts.

4.8 Conclusion

Blockchains are considered as a technological breakthrough to address many healthcare problems for ensuring the protection of sensitive patient's data and preserving their privacy and security. The advantages of decentralizing control and consensus have also motivated the fast adoption of blockchain technology in pharmaceutical supply chains, precision medicine, and many other patient care applications. In this chapter, we introduced the BlockCare framework which proposes an architecture that combines software-defined networking and blockchain to interconnect various IoT elements, SDN networking, and blockchain components to control and manage patients' health-related parameters. Our architecture uses blockchain to bring more transparency and security to the transactions, while SDN enforces scalability and performance by dynamically provisioning blockchain virtual functions for firewalling and mitigating malicious traffic. In light of this architecture, we described the

[3] https://goo.gl/nCQHFB.

different components and algorithms it encompasses. Blockchain is still in its early stages, and soon we should expect an explosive growth of novel solutions that will emphasize providing functional and architectural design approaches to extend its use for ensuring the security of the future Internet, protecting autonomous and self-driving cars, or even for the anonymization of healthcare transactions and obfuscating the patient's identity as it is discussed in the next chapter.

Acknowledgements This work was funded in part by the Eugen Ionescu Visiting Scholars Program and the United States National Science Foundation (NSF) CNS US Ignite 1531079. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Eugen Ionescu Program or NSF.

References

1. Angeletti F, Chatzigiannakis I, Vitaletti A (2017) The role of blockchain and IoT in recruiting participants for digital clinical trials. In: 2017 25th international conference on software, telecommunications and computer networks (SoftCOM), pp 1–5
2. Arias O, Wurm J, Hoang K, Jin Y (2015) Privacy and security in internet of things and wearable devices. IEEE Trans Multi-Scale Comput Syst 1(2):99–109
3. Azaria A, Ekblaw A, Vieira T, Lippman A (2016) Medrec: using blockchain for medical data access and permission management. In: 2016 2nd international conference on open and big data (OBD), pp 25–30
4. Back SA, Corallo M, Dashjr L, Friedenbach M, Maxwell G, Miller A, Poelstra A, Timón J (2014) Enabling blockchain innovations with pegged
5. Bai X, Cheng Z, Duan Z, Hu K (2018) Formal modeling and verification of smart contracts. In: Proceedings of the 2018 7th international conference on software and computer applications, ICSCA 2018, New York. ACM, pp 322–326
6. Bhuiyan MZA, Zaman A, Wang T, Wang G, Tao H, Hassan MM (2018) Blockchain and big data to transform the healthcare. In: Proceedings of the international conference on data processing and applications, ICDPA 2018, pp 62–68
7. Bocek T, Rodrigues BB, Strasser T, Stiller B (2017) Blockchains everywhere—a use-case of blockchains in the pharma supply-chain. In: 2017 IFIP/IEEE symposium on integrated network and service management (IM), pp 772–777
8. Camenisch J, Drijvers M, Dubovitskaya M (2017) Practical UC-secure delegatable credentials with attributes and their application to blockchain. In: Proceedings of the 2017 ACM SIGSAC conference on computer and communications security (CCS'17), pp 683–699
9. Dubovitskaya A, Xu Z, Ryu S, Schumacher M, Wang F (2017) Secure and trustable electronic medical records sharing using blockchain. In: AMIA 2017 annual symposium proceedings
10. E. E. A. Inc. Enterprise Ethereum client specification (2018)
11. Ehmke C, Wessling F, Friedrich CM (2018) Proof-of-property: a lightweight and scalable blockchain protocol. In: Proceedings of the 1st international workshop on emerging trends in software engineering for blockchain, WETSEB'18, pp 48–51
12. Gordon WJ, Catalini C (2018) Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. Comput Struct Biotechnol J 16:224–230
13. Herrera JG, Botero JF (2016) Resource allocation in NFV: a comprehensive survey. IEEE Trans Netw Serv Manage 13(3):518–532
14. IBM Team (2015) Blockchain: the chain of trust and its potential to transform healthcare—our point of view

15. Islam SMR, Kwak D, Kabir MH, Hossain M, Kwak K (2015) The internet of things for health care: a comprehensive survey. IEEE Access 3:678–708
16. Karafiloski E, Mishev A (2017) Blockchain solutions for big data challenges: a literature review. In: IEEE EUROCON 2017—17th international conference on smart technologies, pp 763–768, July 2017
17. Kataoka K, Gangwar S, Podili P (2018) Trust list: internet-wide and distributed IoT traffic management using blockchain and SDN. In: 2018 IEEE 4th world forum on internet of things (WF-IoT), Feb 2018, pp 296–301
18. Khan MA, Salah K (2018) IoT security: review, blockchain solutions, and open challenges. Future Gener Comput Syst 82:395–411
19. Kravitz DW, Cooper J (2017) Securing user identity and transactions symbiotically: IoT meets blockchain. In: 2017 global internet of things summit (GIoTS), pp 1–6
20. Kreutz D, Ramos FMV, Veríssimo PE, Rothenberg CE, Azodolmolky S, Uhlig S (2015) Software-defined networking: a comprehensive survey. Proc IEEE 103(1):14–76
21. Kuo T-T, Kim H-E, Ohno-Machado L (2017) Blockchain distributed ledger technologies for biomedical and health care applications. JAMIA J Am Med Inform Assoc
22. Li C, Zhang LJ (2017) A blockchain based new secure multi-layer network model for internet of things. In: 2017 IEEE international congress on internet of things (ICIOT), pp 33–41, June 2017
23. McKeown N, Anderson T, Balakrishnan H, Parulkar G, Peterson L, Rexford J, Shenker S, Turner J (2008) Openflow: enabling innovation in campus networks. SIGCOMM Comput Commun Rev 38(2)
24. Medicines S (2017) Counterfeit drugs in America: crimes, victims and solutions
25. Mirjalily G, Luo Z (2018) Optimal network function virtualization and service function chaining: a survey. Chin J Electron 27(4):704–717
26. Nehai Z, Piriou P-Y, Daumas F (2018) Model-checking of smart contracts. In: IEEE blockchain 2018—the 2018 IEEE international conference on blockchain, July 2018
27. Peterson K, Deeduvanu R, Kanjamala P, Mayo KB (2016) A blockchain-based approach to health information exchange networks. In: White paper
28. Rajagopalan A, Jagga M, Kumari A, Ali ST (2017) A DDOS prevention scheme for session resumption sea architecture in healthcare IoT. In: 2017 3rd international conference on computational intelligence communication technology (CICT), pp 1–5
29. Salahuddin MA, Al-Fuqaha A, Guizani M, Shuaib K, Sallabi F (2017) Softwarization of internet of things infrastructure for secure and smart healthcare. Computer 50(7):74–79
30. Samaniego M, Deters R (2016) Using blockchain to push software-defined IoT components onto edge hosts. In: Proceedings of the international conference on big data and advanced wireless technologies, BDAW'16, pp 58:1–58:9
31. Shae Z, Tsai JJP (2017) On the design of a blockchain platform for clinical trial and precision medicine. In: 2017 IEEE 37th international conference on distributed computing systems (ICDCS), June 2017, pp 1972–1980
32. Singh M, Singh A, Kim S (2018) Blockchain: a game changer for securing IoT data. In: 2018 IEEE 4th world forum on internet of things (WF-IoT), Feb 2018, pp 51–55
33. Steichen M, Hommes S, State R (2017) Chainguard: a firewall for blockchain applications using SDN with openflow. In: 2017 principles, systems and applications of IP telecommunications (IPTComm), Sept 2017, pp 1–8
34. Urien P (2018) Towards secure elements for trusted transactions in blockchain and blochchain IoT (BIoT) platforms. Invited paper. In: 2018 fourth international conference on mobile and secure services (MobiSecServ), Miami Beach, pp 1–5
35. Vukolic M (2017) Rethinking permissioned blockchains. In: Proceedings of the ACM workshop on blockchain, cryptocurrencies and contracts, BCC'17, pp 3–7
36. Wood T, Ramakrishnan KK, Hwang J, Liu G, Zhang W (2015) Toward a software-based network: integrating software defined networking and network function virtualization. IEEE Netw 29(3):36–41

37. Xia Q, Sifah EB, Asamoah KO, Gao J, Du X, Guizani M (2017) Medshare: trustless medical data sharing among cloud service providers via blockchain. IEEE Access 5:14757–14767
38. Zhang P, Walker MA, White J, Schmidt DC, Lenz G (2017) Metrics for assessing blockchain-based healthcare decentralized apps. In: 2017 IEEE 19th international conference on e-health networking, applications and services (Healthcom), pp 1–4
39. Zhang P, White J, Schmidt DC, Lenz G (2017) Applying software patterns to address interoperability in blockchain-based healthcare apps. In: 24th pattern languages of programming conference, abs/1706.03700
40. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST (2018) FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 16:267–278
41. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. In: 2017 IEEE international congress on big data (BigData congress), pp 557–564

Chapter 5
IoT Performability for Medical Wearable Device by Data Privacy and Fault Tolerance

Raluca Maria Aileni, George Suciu, Carlos Valderrama, and Sever Pasca

Abstract The goal of this book chapter is to set the optimal safety possibility to use wearable biosensors for patient monitoring. This objective consists in minimizing the security risks by using the adequate methodology for working with personal sensitive data, anonymization and data analytics in the cloud. The possibilities of using data are to filter in a private cloud and to anonymize for analyzing in public cloud. The proposed methodology will ensure the security and privacy by using different level of protection for different type of user accounts (medical staff and patients). The anonymization of the patients' information is relevant because data is used for a secondary purpose—predictive analysis modeling for understanding and anticipate the diseases behaviors or for preventive medical actions.

5.1 Introduction

The 5G networks [1] will enable the Internet of Everything (IoE) [2], and this can lead to huge efforts for maintaining the device security and data privacy for users, even if it is very attractive to have many devices interconnected. Statistics show that 50 billion embedded devices (sensors, actuators) will be connected to 5G mobile networks by 2020 [1]. The 5G systems, that are needed to connect more users and receive a huge volume of raw data, are composed of heterogeneous devices with low and high bandwidths. In order to use the entire potential of IoT [3], 5G network must reduce the response times (latency). The sensors and actuators controlled and operated remotely or autonomously in the cloud, without deterministic response

R. M. Aileni (✉) · G. Suciu · S. Pasca
Faculty of Electronics, Telecommunication and Information Technology, Politehnica University of Bucharest, Bucharest, Romania
e-mail: raluca.maria.aileni@gmail.com

G. Suciu
Beia Consult International, Bucharest, Romania

C. Valderrama
Faculty of Engineering, Department of Electronics and Microelectronics, Mons University, Mons, Belgium

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_5

times, will limit the utility of IoT technologies [4]. The 5G waveforms, network densification, massive MIMO (multiple-input and multiple-output), and mm Wave (millimeter wave) communications may be incrementally deployed in time.

Internet of Things (IoT) used for connecting the wearable medical devices (invasive and noninvasive) incurs that a huge volume of data is stored in the cloud, enabling aggregation of data from different devices and different patients [5].

The aspects regarding data protection, security and privacy through IoT are new and represent a challenge for adding a privacy component [6]. The IoT value chain involves a high number of stakeholders (insurance companies, medical institutions, relatives and other companies) and allows a huge volume of data to be collected in the cloud and to be used for predictive analysis. Patients' data can be used to improve the medical act by predictive analytics and for malware companies in order to have a detailed profile of individual patients.

The sharing and data processing through cloud services will lead to increase the storage locations and jurisdictions regarding private data storage. The European Regulation and the Directive regarding personal data privacy, published in the EU Official Journal on May 24, 2016, will give citizens back control over their personal data and will simplify the regulatory environment for digital business. The existing scientific inquiry shows that the main challenges of IoT challenges regarding privacy and data protection for medical use are:

- insufficient control and information asymmetry;
- patients consent (privacy laws are consent-based—if patients give their consent or authorization, the data can be used for other research purposes);
- data fusion;
- data anonymizing for analytics (if the data is anonymized then no consent from patient is required);
- anonymous limitation when patients use services for data aggregation;
- security risks (communication links, storage in a cloud without knowing the exact location);
- privacy by design.

There are warnings of a massive security risk by embedding sensors in everyday devices, such as wearable, as the small size and the limited power processing could inhibit the encryption and security measures. For medical area, the IoT should provide data storage in private cloud, anonymization and data analytics in the public cloud [5, 6]. By using IoT, the benefits are both for patients and medical staff because of the continuous monitoring of the vital signs at home; possibility to register, track and analyze data; improvement of the medical action; increasing the knowledge about diseases. The use of IoT for medical monitoring is important for patients with high-risk factors (elderly, person with disabilities). The security risks of the IoT devices used for medical monitoring are unauthorized access to personal data; launch denial-of-service attack; safety risks (by exploiting security vulnerabilities can lead to physical safety); privacy risks (collection of sensitive personal information).

In 2015, Jia-ji Hou and Kuo-Hui Yeh proposed two secure communication protocols for IoT-based healthcare systems, with single sign-on (SSO)-based authentication scheme and a coexistence proof mechanism for multiple sensors (application field of inpatient safety and medication management) [7]. This book chapter presents a survey regarding security risks and safety measures for wearable biosensors used for healthcare monitoring.

As follows, layout Sect. 5.2 explains about IoT and biomedical wearable devices-evolution, perspectives and challenges. Section 5.3 demonstrates security risks and safety concerns for use of wearable devices. Section 5.4 presents personal data management in private cloud versus public cloud for data analytics. Section 5.5 describes methods for IoT wearable risks reduction by data anonymization. Section 5.6 is about future challenges for wearable IoT technologies for healthcare and lastly conclusions.

5.2 IoT and Biomedical Wearable Devices—Evolution, Perspectives and Challenges

The goal of this book chapter is to set the optimal safety possibility for using wearable biosensors for patient monitoring. This objective consists in minimizing the security risks by using the adequate methodology for working with personal sensitive data, anonymization and data analytics in cloud (Fig. 5.1). The anonymization of the patients' data is important because it allows the legal use of data for predictive

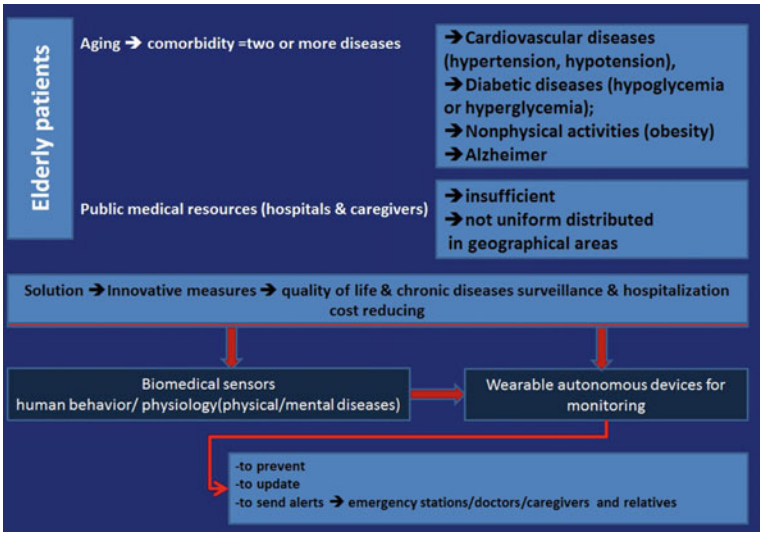


Fig. 5.1 Ecosystem design for patients' health monitoring [8]

modeling, the study of the evolution of the disease and correlations between the various symptoms.

The embedded sensors for a medical purpose will allow remote measurement for blood pressure, temperature, skin moisture, glycaemia or stress and will include actuators to turn on and off devices or make adjustments in real time. Patient vital sign data could be monitored for predictive analytics (Fig. 5.1). Figure 5.1 describes the completely case scenario, why do we need to advance technologically in the field of medicine, when there are aged patients suffering from diseases and medical care center are not enough as required. IoT wearable devices can monitor the patients 24 h in home and in the case of any emergency, alerts are sent to the caretakers as well as to the medical department to take the measures, it is described below. Readings are taken through the biomedical sensors and sent to the wearable autonomous monitoring devices and generate the alerts according to the situation. Understanding the cause and effect of health conditions in correlation with weather parameters (Fig. 5.2), biomedical and environmental conditions are monitor constantly. Figure 5.2 indicates a flow how environmental condition can affect the physiological condition of a patient, how can be monitored, transmitted in the cloud and analyzed repeatedly.

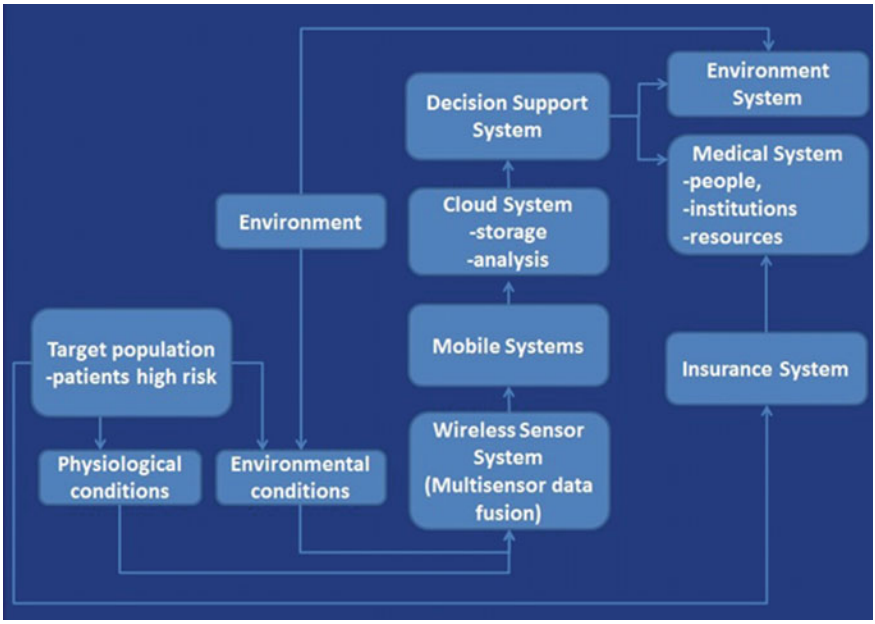


Fig. 5.2 Biomedical and environmental conditions monitoring for patients' risk analysis [9]

5.3 Security Risks and Safety Concerns for Using Wearable Devices

5.3.1 Safety and Privacy Aspects Related to Wearable Medical Devices

Data related to patients can't be released without confidentiality because it contains the most sensitive details of the patient's life, physical or mental health or health history of the patient's family. Individual health data must be complete and updated. This information is used for personal health evaluation, medical prescriptions but also for society in order to improve the medical care and develop predictive models for diseases. Medical data used in statistics must be protected against identity disclosure. According to HIPAA (Health Insurance Portability and Accountability Act) [10], privacy rules accept two standards (Safe Harbor and statistical method) for de-identification of health data. Safe Harbor is stipulating the removing or generalization of 18 variables from health data (e.g. name, city, phone, e-mail, social security number, device identifiers and serial number) [11, 12]. According to data protection in the EU, the patients' sensitive data that identify individuals must not be kept longer than necessary and only when justified by public health important reasons, to derogate from the prohibition on processing sensitive categories of data [13].

5.3.2 Information Privacy for Medical Records

The process of using the data is to filter the information in private cloud and to anonymize for data analysis in a public cloud [14]. The proposed methodology will ensure the security and privacy by using different levels of security for different types of user accounts (medical staff and patients). The anonymization of the patients' data is relevant because the information is useful for a secondary purpose –predictive analyze modeling in order to understand and anticipate the diseases behaviors or for preventive medical actions. Figure 5.3 illustrates data privacy of the patient's data, as seen below.

There are some techniques for data anonymization [15, 16], such as:

- Masking—protect X data by convert in Y data (non-perturbative masking methods, perturbative masking methods and synthetic data generation).
- Synthetic data—protected data X consists of randomly generated records that do not directly derive from the records in Y (fully synthetic data; partially synthetic data and hybrid data).

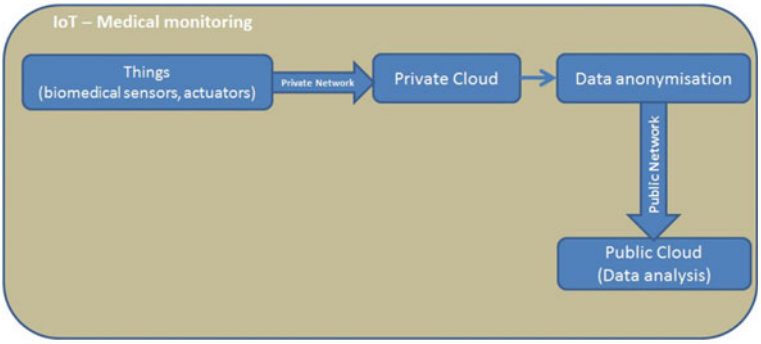


Fig. 5.3 IoT—medical monitoring

For our objective, we choose to use the non-perturbative masking method and synthetic data method with partially synthetic and hybrid data types. For generating the synthetic data, we used the methodology based on three steps:

- set the model for the population
- adjust the model to the original data set X
- generate the synthetic data Y from the model.

The challenges of IoT security for healthcare ecosystem, at the physical and logical level, are cryptographic algorithms development, authentication protocols, access control and privacy [17, 18].

5.4 Personal Data Management in Private Cloud Versus Public Cloud for Data Analytics

Cloud computing architecture for healthcare summarizes several services: applications, Infrastructures as a Service (IaaS), Platforms as a Service (PaaS), and Software as a Service (SaaS) which allow data virtualization, storage and management, as shown below (Fig. 5.4).

The solution for managing personal data saved in private–public hybrid cloud (Figs. 5.5, 5.6 and 5.7). When we have a big volume of data, this is very easy to be used without having local servers for database, but it requires data anonymizing and data process with minimal concern that other user may capture the data in public cloud [19].

Data anonymization in cloud is the process of removing the sensitive information from datasets, in order to anonymize the patients' data. For data analytics in cloud, it is required to anonymize the data from private cloud in order to analyze it in public cloud. For patients' data security in cloud, there are used the anonymization techniques such as *k*-anonymity [20–22], *l*-diversity and *t*-closeness. Figure 5.5 shows

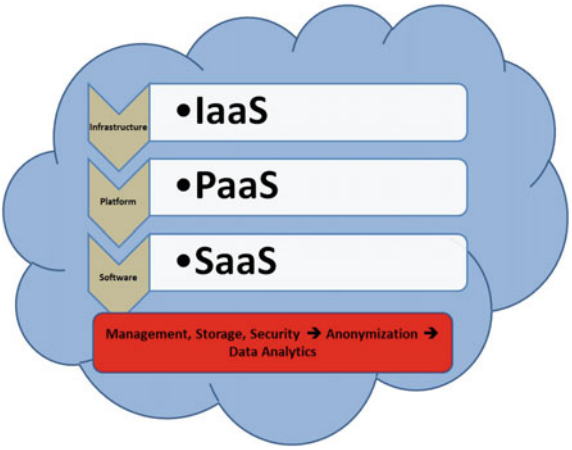


Fig. 5.4 Cloud computing architecture

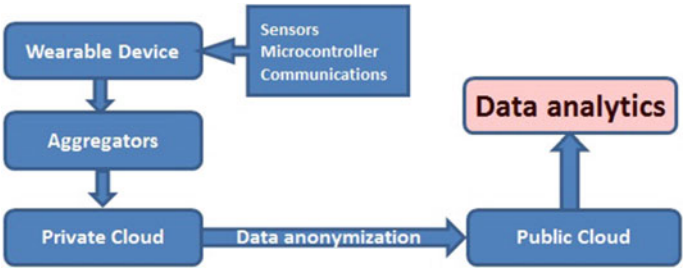


Fig. 5.5 Data management—private—public cloud

that after collecting data from the wearable device (sensors, microcontroller communications), data is aggregated and sent to the private cloud, then data is anonymized before sending it to the public cloud for the sake of patient's privacy and analytics is performed on the public data later on.

We propose a hybrid cloud model based on PaaS architecture (Fig. 5.6) for medical wearable devices. Patients' data should be saved in a private cloud because wearable device involves the usage of wireless personal area network (WPAN) [23, 24].

Starting from the premise that a doctor must have access to sensitive, insensitive, quasi-identifying or identifying attributes of patients' data but patients must not have access to other patients' data or to doctor's data [23], the software solution for data management must be developed for different level of access by setting the permissions rights. Medical data collected from personal medical wearable devices can be used by doctors or for data analytics in order to develop predictive disorders models, but only after data anonymization. Figure 5.7 shows cloud service for health care. As it can be seen, in the first step users (doctor and patient) data is collected from the patients through the biomedical sensors and stored in cloud, doctors can access

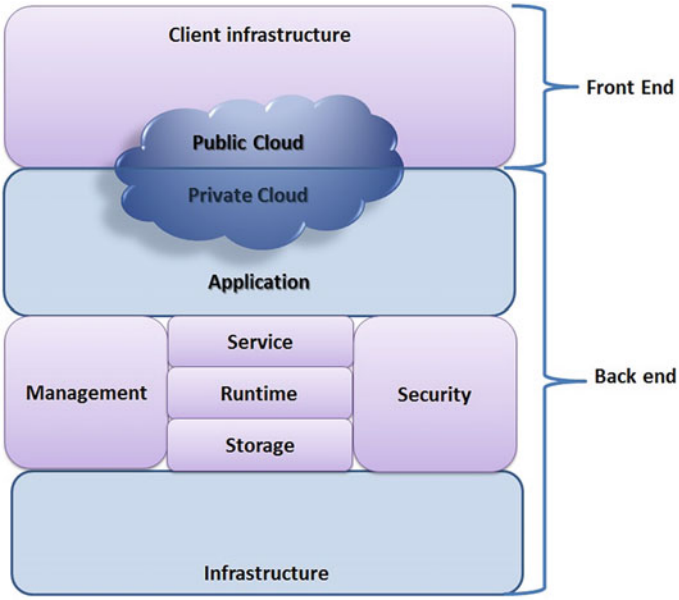


Fig. 5.6 Hybrid cloud PaaS [23]

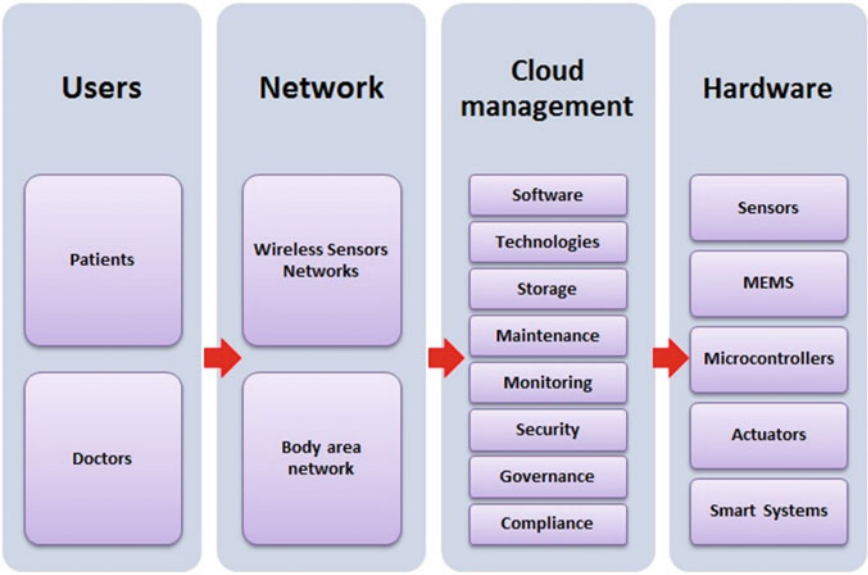


Fig. 5.7 Cloud service for healthcare

the data for patient monitoring purposes then data is aggregated and maintained for future use after making it anonymous, so that patient privacy cannot be leaked.

- *k*-anonymity: The solution of hiding dataset against re-identification by generalizing the attributes that might be used in a linkage attack (quasi-identifiers) [16]. If the attack is targeting a single individual record, then it can be obtained all sensitive attributes [25].

The challenges in *K*-anonymity technique implementation are based on [25] computational complexity of *k*-anonymization algorithm (O(ln *k*)) [16], scalability, robustness and data quality.

The potential attacks are: homogeneity attack (suppose that some sensitive data records contain similar values then the sensitive value for this set of records may be predicted) and background knowledge attack.

- *L*-diversity: This method is an extension of *k*-anonymity that will generate privacy by reducing the granularity of data representation using techniques such as generalization and suppression. *L*-diversity model adds diversity for sensitive values in the anonymization mechanism in order to reduce a potential homogeneity attack.
- *T*-closeness: This method is an extension of *l*-diversity by treating the values of an attribute distinctly by taking into account the distribution of data values for that attribute.

Data anonymization design must take account of the influence factors as environment, medical providers, regulations, applications type, data type, data semantics, utility requirements and privacy [25]. Data privacy by design can be released by anonymization of a few records from a multidimensional data table or by anonymization of a tuple in a record. The objective is to avoid sensitive data disclosure, and this can be achieved by random perturbation methods and group anonymization techniques, such as k-anonymity or l-diversity [25].

5.5 Methods for IoT Wearable Risks Reduction by Data Anonymization

Several methods from statistical disclosure control, as *k*-anonymity, *l*-diversity or *t*-closeness, are required to be used for healthcare data anonymization in order to respect privacy data for patients.

The risks of data anonymization for IoT analytics in cloud are due to insufficient information, missing or incorrect about patients due to l-diversity or l-closeness techniques, or the cohorts are incomplete. In the scientific literature, there are stated two techniques, such as masking and de-identification, in order to remove the link between patient sensitive data and data related to diseases behaviors (age, weight, geographic location, device used, parameters and vital signs values).

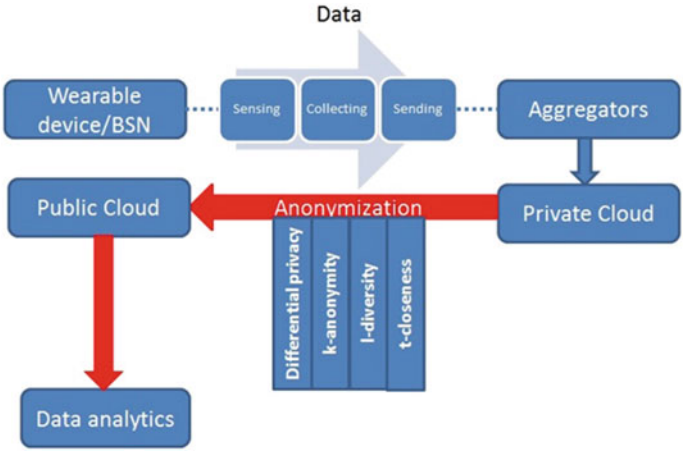


Fig. 5.8 Data flows from analog-to-digital data anonymization [9]

For IoT wearable risks reduction, it is necessary to implement the measures [26] for wearable device authentication and control access and data and privacy protection by anonymization (Fig. 5.8). Figure 5.8 also illustrates data flow from analog-to-digital data anonymization where it can be seen, in the first step, how data is collected from the patients through the biomedical sensors through wearable devices and aggregated to send in private cloud. Four steps are taken while anonymization before making it public for analytics.

5.5.1 Methods Used in Statistical Disclosure Control

The goal of statistical disclosure control (SDC) [27, 28] is to present the linkage and matching of the datasets in order to obtain the sensitive patients' data [29]. For SDC are used perturbative and non-perturbative methods.

Perturbative methods (noise adding, swapping, microaggregation or PRAM) generate a distortion of the original data before publication. This method of masking data is to protect the patients' privacy.

Non-perturbative methods (global recording, sampling or local suppression) generate a partial suppression or a reduction of the original data before publication.

Perturbative and non-perturbative methods can be applied to continuous or discrete data values.

5.5.1.1 Perturbative Methods

Noise addition is a non-perturbative method that can be applied to continuous data. The masking is provided by addition of correlated or uncorrelated noise and linear or nonlinear transformation.

Swapping method can be applied for continuous or discrete data and implies changing the values of sensitive data and maintaining the low-order frequency counts or marginal.

Microaggregation method involves partitioning records, replacing the values in a group with the centroid value of the group and lead to k-anonymity [27].

Post-randomizing method (PRAM) is used only for discrete data. PRAM method implies modification of some discrete variables according to known transition probability matrix [27, 30, 31].

5.5.1.2 Non-perturbative Methods

Generalization (Global recording) method can be applied to continuous and discrete data. For a discrete variable, several categories are combined in order to obtain new categories.

Sampling can be applied for continuous data. The sampling method can be used for sensors, for continuous signal readings and sampling by large time intervals. The time interval for sampling will affect the accuracy of the values recorded (e.g., temperature).

Local suppression method is applied to discrete data and assumes that the values of individual attributes are suppressed in order to increase the set of records accordingly with a combination of quasi-identifier values [32].

5.5.2 Statistical Disclosure Control for Medical Data

Statistical disclosure control is important for statistics in order not to make the sensitive information about patients visible on public clouds [33]. The inconvenience is that by SDC methods, for reducing the disclosure risks, may lead to data loss and generate a data set with information failure. Disclosure type can be patient identity disclosure, attribute disclosure (age, gender, location, disease, wearable devices) and inferential disclosure, based on predictive regression model. The identifiers, such as zip code, birth date or gender, can be used for identifying the patient by related anonymized data with other data (in Tables 5.1 and 5.2, the patients identity linked to diseases, pharmaceutical and wearable devices).

For control of the statistical disclosure in healthcare records are used the following methods:

Table 5.1 Patients' microdata [9]

ID-Name	Zip code	Age	Gender	Disease	Wearable devices
Popescu	112034	67	M	Diabetes	Glucose monitor
Ionescu	112432	75	M	Diabetes	Pulse oximeter, temperature and skin moisture
Ann	203412	58	F	Cardiac Insufficiency	Peacemaker
Mathias	127734	66	M	Arthritis	Pedometer, accelerometer
Clara	112732	77	F	Angina	Wearable ECG
George	231298	72	M	Diabetes	Glucose monitor
Dane	342232	44	M	Angina	Wearable ECG
Carla	331245	42	F	Diabetes	Glucose monitor
Tania	129867	59	F	Angina	Wearable ECG
Marin	324578	66	M	Diabetes	Glucose monitor
Klaus	226743	56	M	Diabetes	Glucose monitor
Flavius	334897	64	M	Arthritis	Pedometer, accelerometer
Dan	376521	73	M	Cardiac Insufficiency	Peacemaker
Adrian	129854	68	M	Angina	Wearable ECG
Flavia	226436	56	F	Diabetes	Glucose monitor
Bianca	296641	67	F	Diabetes	Glucose monitor
Darius	129844	56	M	Arthritis	Pedometer, accelerometer
Adriana	316345	66	F	Angina	Wearable ECG
Popescu	197745	56	M	Cardiac Insufficiency	Peacemaker

Table 5.2 Consumer registration data [9]

Name	Zipcode	Age	Gender	Pharmaceutical product
Popescu	012034	67	M	Amaryl 1mg
Ionescu	012432	75	M	Dialevel
Ann	203412	58	F	Hyperil
Mathias	127734	66	M	Naproxen
Clara	112732	77	F	Nitroglycerin
George	231298	72	M	Dialevel

- randomization method (the process of reducing identification with noise, e.g., fake data name, zip code, city);
- *K*-anonymity method (to reduce the indirect identification from public record database by decreasing the granularity of data representation with the use of generalization and suppression techniques);
- *L*-diversity method (to prevent homogeneity attack in *k*-anonymity by introducing sensitive data diversity);

- *t*-closeness (distance between the distribution of a sensitive attribute in a class and the distribution of the attribute in the whole table is no more than a threshold *t*);
- Differential privacy method.

The disclosure term means re-identification of the patient' microdata (name, phone, age, gender, identity number, address, e-mail) or information about patients based on data released (Tables 5.1 and 5.2).

For anonymization, we used ARX software [34], by setting input data:

- ID-Name attribute as—identifying type, generalization transformation (Fig. 5.9);
- Zip code attribute as quasi-identifying type and generalization transformation (Fig. 5.9);
- Model differential privacy, which allows maximization of the accuracy of queries from statistical databases while the identification of its records, by identifiers and quasi-identifiers, is minimized. For identifiers, we choose ID-name that explicitly identify the record owners [35] and that was removed from the output data released (Fig. 5.10). In Fig. 5.11, there are presented the risks based on quasi-identifiers and is clear that the person can be 100% identified based on ID name and zip code and also by using combinations of 2 IDs (e. q. name/age, name/disease,

Input data | Classification accuracy

	ID -Name	Zipcode	Age	Gender	Disease	Wearable devices
1	Ionescu	112432	75	M	Diabetes	Pulse oximeter, temperature and skin moisture
2	Ann	203412	58	F	Cardiac Insufficiency	Peacemaker
3	Mathias	127734	66	M	Arthritis	Pedometer, accelerometer
4	George	231298	72	M	Diabetes	Glucose monitor
5	Dane	342232	44	M	Angina	Wearable ECG
6	Carla	331245	42	F	Diabetes	Glucose monitor
7	Tania	129867	59	F	Angina	Wearable ECG
8	Flavius	334897	64	M	Arthritis	pedometer, accelerometer
9	Adrian	129854	68	M	Angina	Wearable ECG
10	Flavia	226436	56	F	Diabetes	Glucose monitor
11	Darius	129844	56	M	Arthritis	Pedometer, accelerometer
12	Adriana	316345	66	F	Angina	Wearable ECG
13	Glenn	197745	56	M	Cardiac Insufficiency	Peacemaker

Fig. 5.9 ARX—attribute definition—input data [9]

Output data | Classification accuracy

	ID -Name	Zipcode	Age	Gender	Disease	Wearable devices
1	*	*	75	M	Diabetes	Pulse oximeter, temperature and skin moisture
2	*	*	58	F	Cardiac Insufficiency	Peacemaker
3	*	*	66	M	Arthritis	Pedometer, accelerometer
4	*	*	72	M	Diabetes	Glucose monitor
5	*	*	44	M	Angina	Wearable ECG
6	*	*	42	F	Diabetes	Glucose monitor
7	*	*	59	F	Angina	Wearable ECG
8	*	*	64	M	Arthritis	pedometer, accelerometer
9	*	*	68	M	Angina	Wearable ECG
10	*	*	56	F	Diabetes	Glucose monitor
11	*	*	56	M	Arthritis	Pedometer, accelerometer
12	*	*	66	F	Angina	Wearable ECG
13	*	*	56	M	Cardiac Insufficiency	Peacemaker

Fig. 5.10 ARX—output data [9]

Quasi-identifier	Distinction	Separation
Gender	15.38462%	51.28205%
Disease	30.76923%	79.48718%
Wearable devices	46.15385%	85.89744%
Age	76.92308%	94.87179%
ID -Name	100%	100%
Zipcode	100%	100%
Disease, Wearable devices	46.15385%	85.89744%
Gender, Disease	53.84615%	91.02564%
Gender, Wearable devices	69.23077%	94.87179%
Age, Gender	92.30769%	98.71795%
Age, Disease	100%	100%
Age, Wearable devices	100%	100%
ID -Name, Age	100%	100%
ID -Name, Disease	100%	100%
ID -Name, Gender	100%	100%
ID -Name, Wearable devices	100%	100%
ID -Name, Zipcode	100%	100%
Zipcode, Age	100%	100%
Zipcode, Disease	100%	100%
Zipcode, Gender	100%	100%
Zipcode, Wearable devices	100%	100%
Gender, Disease, Wearable devices	69.23077%	94.87179%
Age, Disease, Wearable devices	100%	100%
Age, Gender, Disease	100%	100%

Fig. 5.11 Risk analysis based on quasi-identifiers [9]

name/gender, name/zip code) or 3 identifiers (age/diseases/wearable device, age/gender/disease).

From Fig. 5.9, in the anonymization tool ARX, the data input is shown for patient participants and their information, such as name, zip code, age, gender, disease and wearable devices. After the operation, the results can be seen in Fig. 5.10, where the important information of the patient is anonymized by providing results, which can be used commonly, taking care of the privacy of the patients.

For quasi-identifier (Fig. 5.11), we used the zip code which could be linked to external data to re-identify individual record owners that was removed from the output data. Re-identification can be achieved by linking and matching two lists (Fig. 5.12).

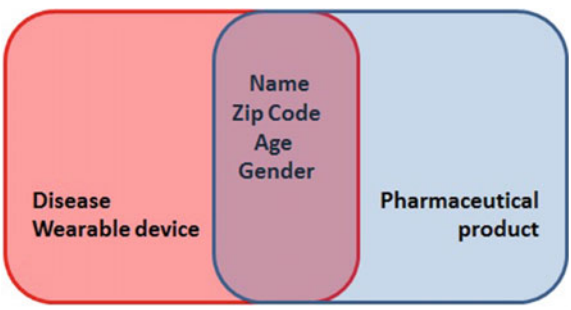


Fig. 5.12 Re-identification by linking and matching

There are three models used for the re-identification attacks are persecutor, journalist, and marketer attack.

In the scenario for the persecutor attack, the intruder has as purpose to re-identify a specific person in a de-identified database. Equivalence classes (unique combinations of quasi-identifiers in the anonymized dataset) measure the re-identification risk [36].

In the second scenario, the journalist attack, the anonymized data is a subset of a larger public database and the journalist is intended to re-identify individuals, not a specific individual. The re-identification risk is calculated and controlled without having access to the larger public database [36].

In the third scenario for the marketer risk, the intruder has as purpose to re-identify as many individuals as possible in a database, even with the risk that the records are misidentified. The marketer risk can be measured by calculating the probability of matching a record in an equivalence class of the de-identified set. After measuring the risk of re-identification, data should be anonymized by de-identification techniques such as record suppression, cell suppression, rounding and aggregation/generalization [36].

By re-identification risk analysis implemented in ARX, we obtained the estimated risk provided for three different attacker models:

- the prosecutor scenario
- the journalist scenario
- the marketer scenario.

In the prosecutor model, the attacker already knows that the data for individual patient is contained in the data set. In the journalist model, the attacker does not know about datasets content. In the marketer model, the attacker is not interested in re-identifying a specific individual but in attacking a larger number of individuals' records.

In Figs. 5.13 and 5.14 are presented the risk analysis for the prosecutor attacker model, journalist attacker model and marketer attacker model, before and after anonymization by differential privacy model.

According to the legal aspects, national laws (Health Insurance Portability and Accountability Act (HIPAA)), privacy rule, and international regulations (the European Directive on Data Protection), personal data privacy is a requirement [10].

The Health Insurance and Portability and Accountability Act from US specify 18 identifiers that must be altered or removed in order to derive a de-identified data set.

The HIPAA privacy rule defines two methods for de-identifying datasets [10]:

- Method based on the removal of a pre-defined set of attributes from the dataset. This method will lead to a reduction of de-identification risk but will also generate data loss.
- Method "expert determination" [37]—statistical disclosure control (SDC) allows balancing privacy risks with information quality (differential privacy or k-anonymity). Using generalization will lead to data precision reduction [38].

Fig. 5.13 Re-identification attack risks before data anonymization [9]

In the biomedical domain, there are used methods for microdata release (k-anonymity) because these involve using non-perturbative methods that allow data correctness [37].

The following types of attributes specified by HIPAA can potentially be detected (Fig. 5.15):

– age
– zip code.

In general, an algorithm for data anonymization is based on adequate transformation selection, data transformation and evaluation of the model quality [39]. Data quality models such as *Non-uniform entropy* or *Height* can be used as objective functions for optimizing the output data (anonymized data) of the anonymization process [39, 40]. The non-uniform entropy model compares the output data with the input data in order to evaluate the loss information, while height model quantifies the loss of information as the sum of the generalization levels applied to all attribute values [39].

In comparison with the transformation, where the balance must lead to increase in privacy with a decrease in data quality [41], the quality of data anonymization must balance the anonymization increase with decrease of data loss [42]. The implementation of the quality models into the ARX de-identification tool for biomedical data

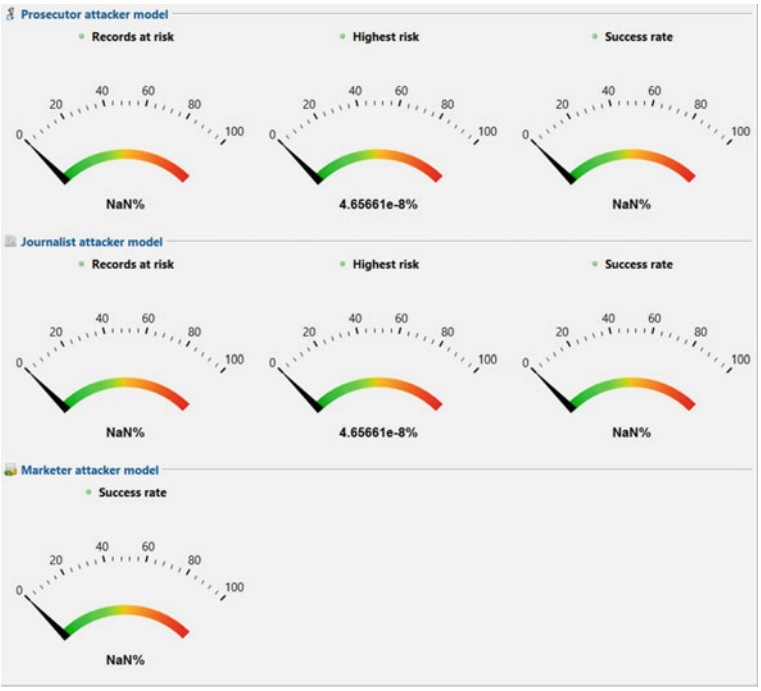


Fig. 5.14 Re-identification attack risks after data anonymization [9]

Column	Identifier	Instance	Match type	Value/Confidence
ID -Name	Name	First name	Attribute value	84.61538%
ID -Name	Name	Last name	Attribute value	61.53846%
Zipcode	Geographic subdivision	ZIP Code	Attribute name	zip code
Age	Date	Date/Time	Attribute name	age
Age	Date	Age	Attribute value	100%

Fig. 5.15 HIPAA identifiers risks [9]

shows that the best-suited model for biomedical application is non-uniform entropy model [42].

5.6 Future Challenges for Wearable IoT Technologies for Healthcare

The wearable medical devices involve IoT scalable applications that can allow data management, security and privacy [43, 44] (Fig. 5.16) with low energy consumption requirements. The challenges for medical wearable devices are the low power consumption, miniaturization, in order to be adopted by healthcare institutions, to

Fig. 5.16 IoT challenges

respond to the European regulation for medical devices, to adapt to the increased and sophisticated needs of the patients, to act as a smart system and must provide safety, reliability, security and sensitive data privacy through software applications.

Because of the huge volume of data, the need to use data for analyzing and for predictive diseases behavior modeling, the wearable medical devices cannot act as only an analogic device and are required to be integrated in smart system (sensors, microcontrollers and communication).

5.7 Conclusions

The purpose of statistical disclosure control used for patients records is data mining for generating new knowledge from data that can be sensitive (personal patients tracking records) [45]. For healthcare data analytics, it is necessary to ensure the patients' data privacy with sensitive data modification by anonymization or removing. The anonymization also has the disadvantage that can lead to insufficient data content or data loss. The proposed method, a hybrid cloud model based on PaaS architecture, takes into account that the software solution for data management should set permission rights and must be architected by a different level of access. For medical data analysis, is often used microdata release (*k*-anonymity) [30] because these involve the usage of non-perturbative methods that allow data correctness [46, 47]. The wearable IoT devices are challenging because of the advantage of biomedical data mining [48] received from sensors and the different security solutions implemented in IoT platforms [49]. Data anonymization involves several methods [46] such as limitations (aggregation), techniques (suppression, swapping, randomization, synthesis, k-anonymity, and differential privacy) [50], transformation (generalization, micro-aggregation), ownership and process guarantees.

As future work, we envision implementing a cloud platform for managing data privacy for wearable medical devices according to GDPR.

Acknowledgements This work was partially supported by UEFISCDI Romania and Ministry of Research and Innovation through project PARFAIT and by European Union's Horizon 2020 research and innovation program under grant agreement No. 787002 (SAFECARE project).

References

1. Maccari L, Karaliopoulos M, Koutsopoulos I, Navarro L, Freitag F, LoCigno R (2018) 5G and the internet of everyone: motivation, enablers, and research agenda. In: European conference on networks and communications (EuCNC). IEEE, pp 429–433
2. Jayaraman PP, Yang X, Yavari A, Georgakopoulos D, Yi X (2017) Privacy preserving Internet of Things: from privacy techniques to a blueprint architecture and efficient implementation. Future Gener Comput Syst 76:540–549
3. Are We Creating An Insecure Internet of Things (IoT)? Security Challenges and Concerns. (n.d.).https://www.toptal.com/it/are-we-creating-an-insecure-internet-of-things. Accessed 10 Aug 2017
4. Murphy KA (2017) Recent FTC regulation of the Internet of Things. Bus Law 72:289–290
5. European Commission. (n.d.) https://ec.europa.eu/justice/data-protection. Accessed 16 July 2017
6. Mahmood Z (2016) Connectivity frameworks for smart devices: the Internet of Things from a distributed computing perspective. Springer, Berlin
7. Hou JL, Yeh KH (2015) Novel authentication schemes for IoT based healthcare systems. Int J Distrib Sens Netw 11(11):183659
8. Aileni RM, Valderrama AC, Strungaru R (2017) Wearable electronics for elderly health monitoring and active living. In: Ambient assisted living and enhanced living environments, pp 247–269
9. Aileni RM, Strungaru R (2019) Theoretical and experimental contributions to the monitoring of vital parameters by using intelligent control systems based on sensors integrated and Cloud Computing service, Ph.D. Thesis
10. Edemekong PF (2018) Health insurance portability and accountability act (HIPAA). https://www.ncbi.nlm.nih.gov/books/NBK500019. Accessed 14 Aug 2017
11. Yadav P, Steinbach M, Kumar V, Simon G (2018) Mining electronic health records (EHRs): a survey. ACM Comput Surveys (CSUR) 50(6):85
12. Garfinkel SL (2015) De-identification of personal information. NISTIR 8053:1–46
13. Art. 9 GDPR—Processing of special categories of personal data (n.d.) https://gdpr-info.eu/art-9-gdpr. Accessed 12 July 2018
14. Veena D, Devidas B (2014) Data anonymization approaches for data sets using map reduce on cloud: a survey. Int J Sci Res 3:308–311
15. Domingo-Ferrer J, Sánchez D, Soria-Comas J (2016) Database anonymization: privacy models, data utility, and microaggregation-based inter-model connections. Synth Lect Inf Secur Priv Trust 8(1):1–136
16. Oganian A, Domingo-Ferrer J (2017) Local synthesis for disclosure limitation that satisfies probabilistic k-anonymity criterion. Trans Data Priv 10(1):61–81
17. Alaba FA, Othman M, Hashem IAT, Alotaibi F (2017) Internet of Things security: a survey. J Network Comput Appl 88:10–28
18. Iyengar A, Kundu A, Pallis G (2018) Healthcare informatics and privacy. IEEE Internet Comput 22(2):29–31
19. Viji KA (2017) A comprehensive cloud security model with enhanced key management, access control and data anonymization features. Int J Commun Networks Inf Secur (IJCNIS) 9(2)
20. Gan VW, Chun-Wei J, Chao H-C, Wang S-L, Yu PS (2018) Privacy preserving utility mining: a survey. In: IEEE 2018 international conference on big data (IEEE BigData)

21. Kavitha S, Sivaraman E, Raja Vadhana P (2014) A suvery on k-anonymity generalization algorithms. Int J Adv Res Comput Commun Eng 2(11)
22. Fei F, Li S, Dai H, Hu C, Dou W, Ni Q (2017) A k-anonymity based schema for location privacy preservation. IEEE Trans Sustain Comput
23. Aileni RM, Pasca S, Valderrama C (2015) Cloud computing for big data from biomedical sensors monitoring, storage and analyze. In: Grid, cloud & high performance computing in science (ROLCG). IEEE, pp 1–4
24. Samarati P, di Vimercati SDC, Murugesan S, Bojanova I (2016) Cloud security: issues and concerns. In: Encyclopedia on cloud computing, pp 207–219
25. Liu X, Xie Q, Wang L (2017) Personalized extended (a, k)-anonymity model for privacy-preserving data publishing. Concurr Comput Pract Experience 29(6):e3886
26. Sun W, Cai Z, Li Y, Liu F, Fang S, Wang G (2018) Security and privacy in the medical Internet of Things: a review. Secur Commun Networks
27. Snoke J, Raab GM, Nowok B, Dibben C, Slavkovic A (2018) General and specific utility measures for synthetic data. J Royal Stat Soc Ser A (Stat Soc) 181(3):663–688
28. Templ M, Meindl B, Kowarik A (2013) Introduction to statistical disclosure control (sdc). Project: relative to the testing of SDC algorithms and provision of practical SDC, data analysis OG
29. Sharma S, Chen K, Sheth A (2018) Toward practical privacy-preserving analytics for IoT and cloud-based healthcare systems. IEEE Internet Comput 22(2):42–51
30. Templ M (2017) Statistical disclosure control for microdata: methods and applications in R. Springer, Berlin
31. Klein MD, Datta GS (2018) Statistical disclosure control via sufficiency under the multiple linear regression model. J Stat Theory Pract 12(1):100–110
32. Data Anonymization: A tutorial—Hamilton Institute (n.d.) https://www.hamilton.ie/privacy2014/josep.pdf. Accessed 10 Oct 2017
33. Shen M, Ma B, Zhu L, Du X, Xu K (2018) Secure phrase search for intelligent processing of encrypted data in cloud-based IoT. IEEE Internet Things J
34. Anonymization tool (n.d.). https://arx.deidentifier.org/anonymization-tool. Accessed 10 Aug 2017
35. Sei Y, Okumura H, Takenouchi T, Ohsuga A (2017) Anonymization of sensitive quasi-identifiers for l-diversity and t-closeness. IEEE Trans Dependable Secure Comput
36. E-Learning Courses on Data Privacy and Anonymization (n.d.) https://www.ehealthinformation.ca/media/e-learning-courses-data-privacy-anonymization. Accessed 20 Aug 2018
37. Prasser F, Eicher J, Bild R, Spengler H, Kuhn KA (2017) A tool for optimizing de-identified health data for use in statistical classification. In: 2017 IEEE 30th international symposium computer-based medical systems (CBMS). IEEE, pp 169–174
38. Prasser F, Kohlmayer F, Spengler H, Kuhn KA (2018) A scalable and pragmatic method for the safe sharing of high-quality health data. IEEE J Biomed Health Inf 22(2):611–622
39. Bild R, Kuhn KA, Prasser F (2018) SafePub: a truthful data anonymization algorithm with strong privacy guarantees. Proc Priv Enhanc Technol 1:67–87
40. Gawande P (2015) SBTAM–balance between privacy and utility using k-anonymity. Spvryan's Int J Eng Sci Techno (SEST) 2(2):18
41. Prasser F, Bild R, Kuhn KA (2016) A generic method for assessing the quality of de-identified health data. MIE, pp 312–316
42. Eicher J, Kuhn KA, Prasser F (2017) An experimental comparison of quality models for health data de-identification. Stud Health Technol Inf 245:704–708
43. Geng H (2017) Internet of Things and data analytics handbook. Wiley
44. Zhou J, Cao Z, Dong X, Vasilakos AV (2017) Security and privacy for cloud-based IoT: challenges. IEEE Commun Mag 55(1):26–33
45. Nethravathi NP, Rao PG, Vaidya CC, Shenoy PD, Venugopal KR, Indiramma M (2017) Generic CBTS: correlation based transformation strategy for privacy preserving data mining. Int J Comput Appl 157(1)

46. Sklavos N (2017) Privacy in a digital, networked world: technologies, implications and solutions. In: Zeadally S, Badra M (eds) Springer International Publishing, p 418
47. Ye Y, Wang L, Han J, Qiu S, Luo F (2017) An anonymization method combining anatomy and permutation for protecting privacy in microdata with multiple sensitive attributes. In: 2017 international conference machine learning and cybernetics (ICMLC). IEEE, 2:404–411
48. Farahani B, Firouzi F, Chang V, Badaroglu M, Constant N, Mankodiya K (2018) Towards fog-driven IoT eHealth: promises and challenges of IoT in medicine and healthcare. Future Gener Comput Syst 78:659–676
49. Arseni SC, Halunga S, Fratu O, Vulpe A, Suciu G (2015) Analysis of the security solutions implemented in current Internet of Things platforms. In: IEEE conference grid, cloud & high performance computing in science (ROLCG), pp 1–4
50. Zeadally S, Badra M (2015) Privacy in a digital, networked world: technologies. Springer, Implications and Solutions

Chapter 6
Toward Trustworthy Cognitive Radio-Based Internet of Medical Things

Jihen Bennaceur, Hanen Idoudi, and Leila Azouz Saidane

Abstract The Internet of Medical Things (IoMT) communication systems are increasingly using wireless networks. However, the use of these networks for such critical systems raised some issues and problems such as the need of allowing more autonomy to the patient with anywhere and anytime monitoring capabilities. Cognitive radio (CR) can be a suitable solution to adapt the wireless network for the IoMT requirements by providing permanent connectivity and, therefore, the spectral availability. In this chapter, we propose a novel design of cognitive radio-based Internet of Medical Things (CR-IMT) networks by providing an efficient, softly and trustworthy integration of cognitive radio technologies in the IoMT applications. Firstly, we propose a cooperative trustworthy spectrum sensing mechanism that enhances the trusty free-band detection accuracy by using a trust and reputation management system. Thus, we formulate two competition noncooperative game models to encourage the cognitive users to trustfully cooperate and to update dynamically the trust values. Moreover, we introduce an opportunistic spectrum scheduling model to ensure the transmission of critical medical data according to the game model outcomes. Extensive simulations validate our approach and prove that it outperforms the traditional established methods in terms of correct detection probability, error probability, data and sensing throughput, delay and residual energy.

6.1 Introduction

The Internet of Medical Things (IoMT) [1] is a worldwide network of interconnected medical objects uniquely addressable. Therefore, this vision of IoMT applications represents a very promising future where billions of medical objects and surrounding environments will be connected and managed through a range of communication technologies and networks. Based on standard communication protocols, it allows the connectivity of people and things at anytime, in anyplace in order to provide a continuous remote patient monitoring in the hospital environment for the vital signs

J. Bennaceur · H. Idoudi (✉) · L. A. Saidane
National School of Computer Science, University of Manouba, Manouba, Tunisia
e-mail: hanen.idoudi@ensi-uma.tn

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_6

such as the heartbeat, the temperature and the blood oxygen. To ensure a continuous connectivity and efficient communications, several technologies are suggested for IoMT environments, from wired to wireless solutions which are gaining more popularity due to their flexibility by being an intrinsic part of modern medical devices. However, the need is raised to mitigate potential issues of electromagnetic interference between the patient's wireless equipment and sensitive medical devices in hospital healthcare environment. Recent trends of research in cognitive radio networks (CRNs) have drawn attention as a potential solution to solve many problems which face the wireless technologies in the IoMT applications such as spectrum scarcity, interference and coverage issues. In this chapter, we introduce a trustworthy integration of CR in the IoMT environment by proposing a hierarchical cognitive emergency application which is composed of very sensitive and critical CR devices. Then, we define a cooperative spectrum sensing basing on the trust and reputation management system in order to ensure a trusty and efficient spectrum performance. Moreover, we consider the problem of dynamic spectrum sharing by formulating two competitive game-based models used to analyze the spectrum sharing situation and the Nash equilibrium is considered as the solution of these noncooperative games which aim at trustfully assigning the CR users according to their reputation values, to encourage the players to cooperate during the spectrum sensing phase and at dynamically and fairly updating the trust system basing on the historical behaviors. Finally, we define an opportunistic scheduling model for our proposal based on the outcomes of the two game designs.

6.2 Cognitive Radio Solution

The latest generation networks used in IoMT applications offer quite satisfactory rates. However, other problems have appeared in these critical applications such as the network connection disparity in some places (inside buildings, remote areas, etc.), the instability of the available data rates and the expensive costs. Moreover, the infrastructure and protocols of these traditional networks have not take into account the constraints of the environment in order to adapt to them in a real time manner. With the very large number of connected objects, the allocation of frequency bands remains costly for the whole amount of transmitted data in e-health applications. As a solution, cognitive radio can be integrated into this application to solve all the problems described above.

6.2.1 Cognitive Radio

Cognitive radio (CR) is next-generation wireless communications, and artificial intelligence has been widely adopted to provide self-learning in order to observe, learn and take action against its operating environment. The concept of cognitive radio was

invented by Mitola, III [2] in 1999 to introduce an intelligent radio that can change its transmitter parameters based on interactions with the environment in which it operates in order to maximize the spectrum use, to increase the throughput and to make the physical layer more reliable. The CR feature enables the unlicensed user [secondary user (SU)] to opportunistically exploit the free channel without interfering with the licensed user [primary user (PU)]. The SU senses periodically the spectrum in order to identify the PU's activities. If the PU is absent, the SU may use the unoccupied spectrum for its current communications. Upon the PU arrival, the SU must leave the channel to avoid the interference with the PU.

6.2.2 Cycle Cognition

To respect the CR concept, the SU should apply the cognition cycle [3] composed by three phases:

1. Spectrum sensing: The CR terminal has sensors to retrieve data on its operating environment. A cognitive radio terminal controls all available spectral bands in a well-defined radio environment, captures their information and then detects unused frequency bands called spectrum holes. The SUs sense the spectrum holes using many enabling algorithms and methods [4] such as the matched-filtering and the energy detector.
2. Spectrum decision: The radio influences its environment through its actuators. The radio can reconfigure itself and modulate this influence according to the situation and QoS objectives. Once the frequencies are detected, the analysis phase spectrum holes will be performed by cognitive nodes to take an accurate decision (using free web/cut another band). The SUs decide individually or cooperatively the channel availability (idle/occupied).
3. Spectrum action: The radio exploits the recovered data to identify the most suitable configuration and to learn from its experiences. First, the electromagnetic radio environment is verified using a CR terminal, according to the following parameters: the transmission rate, the mode of transmission, the transmission bandwidth and the coordination between the mobile cognitive radio terminals. In fact, the decision to use an unoccupied band is related to the occupancy rates and the time of nonuse for the licensed user.

6.2.3 Software-Defined Cognitive Radio (SDCR) Network

Simplifying the resource configuration and the network management is a very challenging and complex task in the conventional networks. Therefore, the software-defined radio (SDR) network has been proposed to implement the cognitive feature, to realize the CRN virtualization and to enhance the spectrum management. SDR [5]

can be an efficient and inexpensive solution to the conventional CRs based on multi-mode, multi-band and multi-functional wireless devices by using software upgrades. In the SDCR environment, most of the components are implemented in the software defined instead of the hardware (modulators/demodulators, amplifiers, etc.) enabling a flexible CR management. Thus, the SDCR concept enhances the CRN performances by allowing the SDCR devices to tune into many different frequency bands simultaneously and assuring multiple communications with other users.

6.3 Motivations for Using CR in IoMT

Nowadays, many motivations are discussed for the integrating of cognitive radio technologies in IoMT communications [6]. Several challenges in the IoMT networks can be efficiently addressed through the CR features such as those discussed hereafter.

6.3.1 Communication Range

Day by day, the communication in the medical environment (between medical devices and cloud servers) depends on the wireless techniques. However, the use of wireless solutions, such as Bluetooth and Zigbee, has many potential limits like the limited range which can decrease the network performance in such critical environment. The CR offers a higher communication range to the IoMT applications, thus paving the way for CR-based IoMT paradigm [7].

6.3.2 Spectrum Scarcity

It presents the main motivation which comes from bandwidth allocation for IoMT-connected objects. Recently, IoMT-connected objects are massively growing in huge numbers making very difficult to allocate the bands to these medical objects. According to Ericsson, 50 billion connected devices will exist by 2020 in the very near future which generates a new challenge for existing communication networks in terms of spectrum availability and band congestion. Thus, in the hospital environment with connected medical devices, this challenge can be more crucial due to the urgent medical data to be transmitted. Traditional communication such as wireless technology cannot support spectrum sharing among multiple medical devices. With dynamic opportunistic spectrum access capabilities offered by the cognitive radio technology, existing spectrum can be utilized more efficiently in order to avoid the potential shortage of spectrum and to support large-scale data transmission in the crucial medical environment.

6.3.3 Interference

In IoTM environment such as emergency framework, the multitude of connected medical devices aiming at operating by using the limited band's resources, arise the interference issues between users. Therefore, CR-based IoMT networks can avoid and alleviate interference limitations by looking for interference-free channels through dynamic spectrum access capability by applying the cycle cognition features. The process of identifying the interference is an indispensable goal of spectrum detection starting with providing the state of the spectrum (free/busy), so that the spectrum can be accessed by a secondary user under the constraint of interference.

6.3.4 Connectivity Issues

In a dynamic environment, users (primary/secondary) can frequently transit their communications to another available band by changing their operating frequency. Due to the mobility of spectrum, cognitive radio devices can dynamically exploit unoccupied channels by allowing radio terminals to exploit the best available frequency (e.g., the frequency with the highest QoS), to maintain the communication requirements transparent during the move to a better band. However, in hospital applications, the doctors, the nurses and the patients are frequently mobile. Mobility will be an indispensable aspect for the future of IoMT structures, and it will be very difficult to ensure the continuity of communication everywhere and anytime. As a solution, medical devices equipped with CR capability can achieve seamless connectivity.

6.3.5 Energetic Requirements

Optimizing the energy consumption becomes an important need to enhance the network lifetime in the IoMT applications. For example, to repair the battery of a heart pacemaker device (a small device placed in the chest or abdomen to help abnormal heart rhythms basing on electrical pulses to prompt the heart to beat at a normal rate) surgery is required which can be very critical for the elder health. A fundamental requirement in IoMT communication is energy efficiency. As a solution, the use of cognitive radio can be considered as green communication for IoMT applications due to the possibility of the adaptive adjustment of the transmission power levels based on operating environments.

6.4 Challenges of the Integration of CR in the IoMT

The integration of CR technologies into the IoMT applications faces critical issues and challenges that need to be addressed prior to deploying CR in medical environments.

6.4.1 Sensing Capabilities

The detection and the classification of the signals is the most indispensable challenge in order to exploit the unoccupied bands. This task is achieved by the cycle cognition process. Cognitive radios added to the medical devises must be able to distinguish signals of licensed user and unlicensed user. However, medical care environments such as hospitals are characterized by the presence of multiple licensed users and a variety of medical signals (oximetry (OXI), blood pressure (ART), respiratory rate (RR), ECG data, etc.) which requires additional solutions. In the literature, many researchers are studying the problem of signal detection in order to find the best method to detect and to classify signals even in a noisy environment. In [8], the authors proposed a CR-based platform for IoMT applications which can protect the medical sensors from harmful interference by adapting the transmit power of wireless devices based on EMI constraints.

6.4.2 Software Radio Issues

Cognitive radio is defined as a technology that allows, using a software radio, to define or modify the operating parameters of the radio frequency of a node. The evolution of software radio is leaded to the apparition of cognitive radio concept with the same issues of the software radio such as improving the data converter process, the agility and enhancing the flexibility of frequencies. The consideration of software radio issues in the CR-based IoMT can improve the network performance by leading to efficient opportunistic spectrum utilization. However, no standard mechanisms have yet been proposed in the literature in order to validate the CR concept and implementations which is still an open door to the future research.

6.4.3 Hardware Issues

The new cognitive feature added to the medical devices is subject to the hardware design evolution. The need of an efficient and an appropriate design, which is capable to sense and evaluate the available frequencies, grows day by day with the increasing

number of the integrated CR-IoMT devices in the medical environment. The new design must address certain hardware requirements such as the antenna sizes and range, and the transmission and the reception power rates.

6.4.4 Security Issues

Cognitive radio networks are vulnerable to conventional wireless network attacks [9] but suffer further from specific attacks [10] targeting the cognitive features. In the network, the cognitive devices, competing to exploit the unoccupied bands, can be honest, selfish, faulty or malicious. Faulty users may send incorrect sensing outcomes due to defective device, fading location or shadowing zone. The attackers aim at monopolizing selfishly the white space causing the interference with the other devices.

These attacks can have more harmful effect in a critical real-time environment such as the hospital areas. The proposed security mechanisms face many issues when applied to CRNs (memory and computation costs, power consumption, etc.). Wireless standards have different security strategies. For example, in WLAN and personal area network (PAN), the only mechanism to incorporate security is identity authentication. The trust and reputation management (TRM) can be the solution for the security issues, as it is an emerging technique to secure the medical environment based on autonomous honest nodes [11].

6.5 Backgrounds

A new e-health design based on the CR infrastructure is proposed in [12]. Thus, the authors solved some of the technical challenges in the design of such systems, such as location sensing, interference avoidance and spectrum allocation among devices.

The authors in [13] introduce a novel cognitive-radio-based approach in order to solve the wireless communication issues for IoMT applications in a hospital environment. Thus, the proposed CR system is designed for the IoMT environment to protect the medical devices from harmful interference by making an adapted transmission power of wireless devices based on electromagnetic interference constraints. In this work, the authors focused on defining a channel access mechanism in the hospital application to determine the transmission parameters for the cognitive devices to avoid the interference. Moreover, the authors of [14] proposed a CR-enabled hospital called Cogmed where the medical devices are implemented by CR features and are classified according to medical urgency and location of hospital. Thus, a queuing theory-based scheduling is introduced for the Cogmed approach based on a priority mechanism where medical and non-medical devices are scheduled according to their functions and locations in the hospital in order to ensure a robust scheduling of

medical data transmission. In [15], the authors discussed the integration of CR technology in a hospital environment with sensitive medical devices as primary users (Pus) and the patient's mobile equipments as the CRs. To mitigate the interference, the authors compute the overlap region to characterize the CR mobile node coverage areas. The estimation of the positions of the patients with CR equipments based on the "gray model" theory allows decision-making to mitigate radio interference. Similar work was done in [16] by proposing a novel architecture of a wireless communication link between the ambulance and the hospital based on CRNs. The proposed cognitive radio module allows the suitable empty frequency band after the spectrum sensing phase on the available channels.

Indeed, the users (sender/receiver) in the ambulance efficiently measure the wideband spectrum and search for the white space to send the data. This work mitigates the interference effect and highlights the CRN performance in patient monitoring.

Furthermore, in [17], the authors proved the ability of cognitive radio to enhance wireless body area network (WBAN) performances. In fact, the exposed scheme adaptively selects the optimal hop number to the sink and adjusts the redundancy in order to minimize the expected latency and maximize the energy saving. In order to facilitate interference detection and mitigation, the authors in [18] define a new protocol to enable WBAN operation within the IoT framework. They promote the integration of a Bluetooth Low Energy (BLE) technology transceiver and a cognitive radio (CR) module within the WBAN coordinator.

However, the CR-based IoMT approaches described above focus only on the interference mitigation and detection issues. Moreover, the most used approaches in the literature do not consider the spectrum sensing phase in hospital environment. All the recent research concentrates on the channel assignment phase. Thus, most of the authors built their schemes according to classical methods and they disregard the privilege of using an intelligent and learning scheme to solve the interference constraints such as game theory models. The authors rarely treated the problem of the critical medical data transmission. In fact, with the CR technology, some CR nodes can be waiting for an unoccupied channel allocation. However, due to the sensitive nature of hospital environment, the critical measures and urgent data must be transmitted in real time without any waiting time. For example, the insulin injection is an urgent data which must be treated and transmitted before the other data without any delay. At this end, we propose an exhaustive cognitive radio-based IoMT framework in order to trustworthily integrate the CR technologies in the hospital environment by mitigating the interference. We introduce two competitive game models to our proposal in order to enhance the spectrum sensing and scheduling phases.

6.6 Trustworthy Proposed Mechanism

In this section, we will introduce the integration of the CR feature in the IoMT applications by defining the description and principals of our proposed framework.

6.6.1 Description and Principals

Our proposed application aims at monitoring and treating the patients in the emergency environment. It consists of a set of interconnected devices or sensors worn by or implanted in the human body in order to collect critical medical information and to interact directly with the human system. When a sensor detects a serious abnormality (heart attack and blood clot), it sends an emergency message to the hospital to claim an ambulance. In the emergency, the patients are monitored by the Data Fusion Center (DFC) which collects the vital signals from all patients. The monitored vital signals can be: temperature (T), oximetry (OXI), blood pressure (ART), respiratory rate (RR), ECG data, etc. Each signal is collected by a dedicated cognitive sensor. The patient's vital signals are analyzed and processed by DFC. Depending on the symptoms analyzed on each patient, the doctor/patient adjusts the active sensor parameters. For example, if a patient suffers from high blood pressure, the doctor should adjust the dose of drugs injected by the active pressure sensors. We can classify the healthcare information into two categories:

- Urgent medical data: It is the information collected from the active cognitive sensors which must be transmitted to the DFC without any delay such as the insulin injection request.
- Routine medical data: It is less priority than the urgent data. This data is resulted from the passive cognitive sensor periodic measurement of body vital signals such as the temperature.

6.6.2 Hierarchical Architecture

Our proposed framework consists of a large number of low-power sensors that can perform detection, simple calculation and short-distance communication. The cognitive sensors implanted in or on the patient body are battery powered and have a predefined lifetime. However, the long-distance transmission can be very expensive since the energy consumption is proportional to the transmission distance. Therefore, we propose the deployment of relay nodes [19] which are more powerful nodes. Their main task is to facilitate the communication with sensors and with other relay nodes to ensure network connectivity and optimize energy depletion (due to data transmission). For emergency application, monitoring and treatment of the patient, our proposed infrastructure (see Fig. 6.1) forms a hierarchical architecture composed by the DFC, the node relay and the cognitive sensors (active/passive). The zone is divided into N_{cell} grids, and a relay node is placed in each cell. The cell number is dependent on the relay node coverage which is expressed by the following formula:

$$N_{\text{cell}} = \frac{\text{Surf}}{R_{\text{relay}}} \tag{6.1}$$

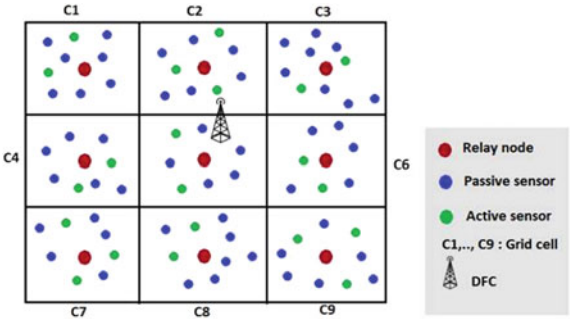


Fig. 6.1 Proposed hierarchical architecture

where the Surf is the emergency surface and R_{relay} is the relay node range.

The emergency of medical data transmission and the need to have robust communication architecture require an adequate topology that is able to provide a low-cost communication links for ensuring reliable connectivity anytime and anywhere. Our proposed hierarchical architecture can be an appropriate key for these challenges which can:

1. Ensure a high connectivity ratio compared with the random cognitive sensor deployment.
2. Guarantee the reliability and scalability for large-scale CRIMT networks.
3. Extend network lifetime in terms of connectivity and coverage by using the relay nodes which minimize the energy consumption.
4. Deliver green communication for the CRIMT applications.

6.6.3 Assumptions

- We consider a cognitive radio-based Internet of Medical Things (CRIMT) network composed of N_S SUs and N_P PUs intended to be distributed over a given region.
- All nodes are heterogeneous: They differ in terms of function characteristics (for instance: temperature detection, insulin injection, heartbeat monitoring, etc.) and the data transmission priority.
- The nodes are placed into a grid topology. In each cell, a relay node is statistically placed in order to maximize the energy gain and DFC is placed in the middle of the covered field.
- We propose a realistic real-time multi-channel environment with N_P PUs.
- The CRIMT is characterized by a centralized cooperative detection where the SUs periodically send their local detections to the DFC in each sensing timeframe to make the final detection about the channel availability.
- The devices are classified into two main categories: Active cognitive (AC) device acts directly on the patient and on the other devices (e.g., the insulin injection sensors, sensors, ECG, etc.), and passive cognitive (PC) device performs the

control and monitoring of patients such as the incubators, infusion pumps and anesthesia machines.

- We suppose that the AC has the highest energy level and the highest priority to use the free spectrum compared with the PC with the lowest spectrum allocation priority.
- When the ACs require the free channel to transmit their critical urgent messages and there is no available spectrum, the ACs must reduce their transmission signal in order to use the occupied channel without interference with the PUs.

6.6.4 Cooperative Trustworthy Spectrum Sensing

In this section, we will illustrate our proposed design based on the trust management and reputation mechanism by introducing trustworthy spectrum sensing scheme in order to define a robust communication architecture. In the cognitive cycle, the spectrum sensing presents the most critical task to detect the availabilities of bands. To this end, our scheme exploits the cooperative spectrum sensing interest in order to enhance better the spectrum sensing performance by making an accurate decision. In each spectrum sensing phase, the cognitive sensor (active/passive) senses periodically the channels to make a local decision $D_i(t)$ about the PU activities. Then, DFC gathers the spectrum sensing from all the cognitive users weighted by their reputation values to make the final decision denoted by AG based on the aggregation function explained by the following formula.

$$\mathrm{AG}(t) = \frac{\sum_{i=1}^{N} D_i(t) * R_i}{\sum_{i=1}^{N} R_i} \tag{6.2}$$

The DFC assigns a reputation value for each cognitive sensor depending on its historical trusty detection in order to make an accurate and robust decision.

After the data aggregation step, the DFC compares gathered final decision AG with each detection of the users. Each mismatch detected signifies the presence of faulty detection of the cognitive node which requires the intervention of DFC by updating its reputation value. Furthermore, the correct detection of the CR node, which is defined as the correspondence between the $D_i(t)$ and the $\mathrm{AG}(t)$, is rewarded by increasing the reputation values of the trustworthy user. The new reputation values help to compute correctly the final decision which improves the spectrum sensing performance.

The reputation update process can be formulated according to the following mathematical expression:

$$R_i(i) = R_i(i-1) + (-1)^{\mathrm{AG}(t)+D_i(t)} \tag{6.3}$$

6.7 Dynamic Multi-winners' Spectrum Access Games

In this section, we model the different interactions between the communicating cognitive nodes using two competitive game models which are distinguished by the participated players, the adopted strategies and game cooperation aspect. We propose two game designs for our model in order to:

- Dynamically update the reputation value basing on learning method included with the game theory concept.
- Define a trustworthy channel assignment mechanism by introducing the trust concept as a scheduling order priority.

We formulate the different interactions as game models denoted as $\mathrm{Ga}^i = \left[\mathrm{Na}, A_j^i, \pi_j^i\right]$ where $i \in [1, 2]$ and $\in [1, N_s]$.

During the game process, the player aims to maximize its profit individually without cooperation with each other.

- The DFC is the arbitrator which manages the game progress, sorts the winning players' order and updates the reputation values.
- A_i^1 are the static strategies adopted by the two players. The strategies for both the players are "Sends correct local detection" and "Sends a faulty local detection."
- π_i^1 defines the payoff of utility function used to define the gain of each player during the cooperative spectrum sensing phase.

6.7.1 Competitive Game Models

The two proposed games are characterized by the competitions between the players. Every cognitive user is mainly interested to maximize its personnel payoff, and therefore all the decisions are made competitively and moreover selfishly.

6.7.1.1 First Proposed Game Model

The first game model illustrates the interaction between the active cognitive sensors competes to use the unoccupied channels. In order to identify the channel assignment order for the active users, a noncooperative simultaneous finite game denoted as $\mathrm{Ga}^1 = \left[\mathrm{Na}, A_i^1, \pi_i^1\right]$ is introduced to our scheme. The AC_i and the AC_j are the two players participating in our game model. Table 6.1 resumes the payoff matrix of our game model.

Table 6.1 Payoff matrix for the AC_i and the AC_j

	AC_i sends correct detection	AC_i sends faulty detection
AC_j sends correct detection	$U_t(AG(t), D_i(t), R_i(t)) - C_d$,$U_t\big(AG(t), D_j(t), R_j(t)\big) - C_d$	$U_t(AG(t), D_i(t), R_i(t)) - C_d$,$- U_f\big(AG(t), D_j(t), R_j(t)\big) - C_d$
AC_j sends faulty detection	$-U_f(AG(t), D_i(t), R_i(t)) - C_d$,$U_t\big(AG(t), D_j(t), R_j(t)\big) - C_d$	$U_f(AG(t), D_i(t), R_i(t)) - C_d$,$U_f\big(AG(t), D_j(t), R_j(t)\big) - C_d$

Table 6.2 Payoff matrix for the PC_i and the PC_j

	PC_i sends correct detection	PC_i sends faulty detection
PC_j sends correct detection	$U_t(AG(t), D_i(t), R_i(t)) - C_d$,$U_t\big(AG(t), D_j(t), R_j(t)\big) - C_d$	$U_t(AG(t), D_i(t), R_i(t)) - C_d$,$- U_f\big(AG(t), D_j(t), R_j(t)\big) - C_d$
PC_j sends faulty detection	$-U_f(AG(t), D_i(t), R_i(t)) - C_d$,$U_t\big(AG(t), D_j(t), R_j(t)\big) - C_d$	$U_f(AG(t), D_i(t), R_i(t)) - C_d$,$U_f\big(AG(t), D_j(t), R_j(t)\big) - C_d$

6.7.2 Second Proposed Game Model

After the first game model progress, we aim at defining the interactions between the passive cognitive sensors as second game model where the users intend to win individually the access to the channels. This game is proposed to our scheme as a noncooperative simultaneous finite game denoted as $Ga^2 = \big[Na, A_i^2, \pi_i^2\big]$. The PC_i and the PC_j are the two players participating in our game model. Table 6.2 resumes the payoff matrix of our game model.

6.7.3 Game Model Analysis

Ga^i is a novel multi-winner spectrum assignment game model. In fact, allowing multiple winners to lease the available bands is the core of our proposal. Indeed, the purpose of our game model is to assign idle bands to the players according to their reputation values. Indeed, the arbitrator DFC compares the final decision $AG(t)$ with the local decision of each player to decide the scheduling order. During the game progress, the DFC sorts the nodes by their reputation values in order to treat fairly the players and to assign trustfully the channels. The player with the highest trust values allocates the first available channel. The second allocates the first channels if no interference has been detected with the previous allocated nodes. All the winners repeat the same assignment process using a decreasing order. Thus, the game model

is used by our proposal in order to update the reputation values of the players based on the historical spectrum detections during the cooperative spectrum sensing phase.

6.7.4 Nash Equilibrium Analysis

The comprehensive benefit is $U_t - C_d$ when secondary users (active/passive) adopt "sending correct detection" strategy regardless the strategy of its adversary where the C_d is the energetic cost of spectrum sensing. When one of players chooses to play selfishly by sending a faulty detection, the DFC punishes this player with a loss equal to $-U_f$. However, when the two players incorporate together to trick the DFC by making an identical strategy, both have profits equal to $-U_f$.

Our two game models $\mathrm{Ga}^i = \left[\mathrm{Na}, A_j^i, \pi_j^i\right]$ have two Nash equilibriums. The first one is the combination of strategy "sending correct detection" with the point couples ($U_t - C_d$,$U_t - C_d$). It is an equilibrium since the players prefer to send correct detection during the spectrum sensing which can be rewarded by exploiting the idle channel. However, the second Nash equilibrium is the combination of strategy "sending faulty detection" with the point couples (U_f, U_f). In this case, the players aim at depriving other users of the free channel access by selfishly monopolizing the white space. We begin with the following lemma to analyze the Nash equilibrium:

Lemma 1 *A strategy profile* $s^* = \left(s_1^*, s_2^*.., s_n^*\right)$ *is a pure strategy Nash equilibrium (PSNE) if, for every player i and any other strategy* $s_i^{'}$ *that player i could choose,* $u_i\left(s_i^* s_{-i}^*\right) \geq u_i\left(s_i^{'} s_{-i}^*\right)$.

This game has two pure strategy Nash equilibriums which can be represented formally as:

$$\mathrm{PSNE} = \left\{(U_t - C_d, U_t - C_d), \left(U_f, U_f\right)\right\} \tag{6.4}$$

In cognitive radio networks, secondary users always have a tendency to use selfish strategy in order to obtain more benefits. The selfish behavior among secondary users will make the cooperative spectrum sensing accuracy lower, and reduce each cognitive user's own benefits. At this end, we aim at computing the optimal probability for each user to make a correct detection basing on the mixed strategy concept.

Lemma 2 *Mixed strategy is a set of probability measures associated with players. Let* S_i *be the set of possible actions of player.* $\forall s_i^j \in S_i$, P_j *denotes the probability of performing action* s_i^j. *We have*:

$$\sum_{s_i^j \in S_i} P_j = 1 \tag{6.5}$$

According to the two previous lemmas, we attribute a probability for each strategy adopted by the two players:

$$\text{IP(Player } i \text{ sends a correct detection)} = p \tag{6.6}$$

$$\text{IP(Player } i \text{ sends a correct detection)} = 1 - p \tag{6.7}$$

To find mixed strategies, we aim at computing the probability p:

$$p * (U_t - C_d) + (1 - p) * (U_t - C_d) = p * -U_f + (1 - p) * U_f \tag{6.8}$$

After simplification, we derive the following equation:

$$p = \frac{U_f - U_t + C_d}{2 * U_f} \tag{6.9}$$

We conclude that players adopt the cooperative strategy with the following condition:

$$p > \frac{U_f - U_t + C_d}{2 * U_f} \tag{6.10}$$

To discourage the cognitive nodes from being selfish by inciting the cooperative behaviors, our model must satisfy the previous condition.

6.8 Opportunistic Spectrum Scheduling Scheme

Considering the limited sensing capability, the complexity of the emergency network and the crucial medical data, our aim is to carefully coordinate and schedule the SUs, in order to maximize the spectrum opportunities, while protecting PUs from interference. Boosted by the promising opportunistic assignments proposed in [20] for the CRNs, we will introduce our opportunistic scheduling mechanism based on the two game models Ga_i.

6.8.1 Scheduling Principals and Mechanism

We propose a centralized scheduling executed by the DFC based on the following principles:

1. The DFC sorts the SUs based on the priority indicating the order according to which the SUs are scheduled.

2. ACs as insulin injection sensors have the highest priority to use the channel than passive nodes.
3. The two main priorities assigned by the DFC are:
   (a) $\boldsymbol{P}_{\mathrm{A}}$ **priority** is the priority assigned to the active nodes.
   (b) $\boldsymbol{P}_{\mathrm{P}}$ **priority** defined the passive node priority where $\boldsymbol{P}_{\mathrm{A}} > \boldsymbol{P}_{\mathrm{P}}$.
4. In order to establish an effective channel assignment order, DFC executes firstly the game model Ga_1 for users having priorities P_{A}. Secondly, it executes the game Ga_2 for the passive nodes having priorities P_{P}.
5. The DFC schedules the nodes having the same reputation values by using the sensor identifier (ID).
6. For any time slot and on a given channel C_{I}, DFC schedules any node that does not interfere with the current scheduled SUs and there are no primary users transmitting on this channel.
7. Before the data transmission permission, DFC must check the interference constraints [21] which can be classified into:
   (a) Inter-cell interference: The DFC must assign the channels while checking the absence of interference between the SUs of the same cell.
   (b) Intra-cell interference: The DFC must allocate the available channels while ensuring the absence of interference between the SUs of neighboring cells.

However, the ACs scheduling cannot lead necessarily to the channel assignment of all active nodes in the CRSN. Due to the importance of critical data to be transmitted, we introduce a solution which allows the active nodes to use the occupied channel without interfering with the SUs.

6.8.2 Scheduling Algorithm

We introduce the new scheduling algorithm for our proposal which is executed by the DFC in two main steps:

1. Firstly, the algorithm 1 determines the set of ACs denoted as $\mathrm{Set}_{\mathrm{A}}$ scheduled in the time slot $t = 1$. All active nodes must be scheduled for the first time slot without any waiting time. Thus, the DFC sorts the set $\mathrm{Set}_{\mathrm{A}}$ based on the game Ga_1 outcomes. Moreover, if two nodes have the same reputation values then the node with smallest ID number is firstly scheduled. The AC which does not conflict with the current scheduled nodes allocated the same channel. In contrast, to avoid interference problems, any other conflicting active node in the sorted set $\mathrm{Set}_{\mathrm{A}}$ is assigned to a different available channel. Sometimes, the DFC cannot assign channels to ACs because of interference constraints or because the channels are no longer free. As a solution, we propose that the active sensor reduces its transmission signal so that it can transmit the critical medical data without interfering with the SUs occupying the free channel.

2. Secondly, the DFC sorts the set of PCs scheduled in the current time slot t designed by Set_P based on the game Ga_2 outcomes. Thus, two equal nodes in terms of reputations are sorted according to their ID numbers in ascending order. Based on algorithm 2, the DFC repeats the same channel assignment concept for the Ga_2. However, the passive nodes without allocated channel may wait for the channel vacancy in the next time slot $t + 1$ without reducing the signal transmission.

The algorithm ends when all nodes (active/passive) have allocated by the number of slots required.

Algorithm 1 Active Node Scheduling

```
/* Schedule active SUs */
Output: N_C= list of available channels. N_A=List of active SUs having packets to transmit.
Output: Schedule the active SUs
Set_A ← active sensors with priority P_A
Sort (Set_A) according to the game Ga^1
Sort the nodes having the same reputation values according to their ID
Initialize the set of nodes scheduled on the current channel C denoted as List_C
α ← First node in Set_A
while α do
    α ← Next node in Set_A
    /* try to schedule α by using the function GeneralSchedule(α, t, nb_channel, List_C) with t=1 (all active nodes must be scheduled for the first time slot without waiting time) */
    (test,channel)=GeneralSchedule(α, 1, nb_channel, List_C)
    if test==True then
        The node α has been scheduled
        N_A ← N_A\{α}
    else
        The node α cannot be scheduled
        Set of nodes without channel assignment
        Set_A* ← α
    end if
    α ← Next node in Set_A
end while
α* ← First node in Set_A*
while α* do
    α* ← Next node in Set_A*
    Reduce signal transmission of α*
    Channel assignment for α*
    /*The node α* has been scheduled*/
    N*_A ← N*_A\{α*}
end while
```

Algorithm 2 Passive Node Scheduling

Output: N_C= list of available channels. N_P=List of passive SUs having packets to transmit.
Output: Schedule the passive SUs
$Set_{P'} \leftarrow$ active sensors with priority $P_{P'}$
Sort (Set_P) according to the game Ga^2
Sort the nodes having the same reputation values according to their ID
Initialize the set of nodes scheduled on the current channel C denoted as $List_C$
$\beta \leftarrow$ First node in Set_P
while β **do**
  $\beta \leftarrow$ Next node in Set_P
  /* try to schedule β by using the function GeneralSchedule*/
  (test,channel)=GeneralSchedule($\beta, t, nb_{channel}, List_C$)
  if test==True **then**
    The node β has been scheduled
    $N_P \leftarrow N_P \backslash \{\beta\}$
  end if
  $\beta \leftarrow$ Next node in Set_P
  $t \leftarrow t + 1$ //Next slot
end while

Algorithm 3 GeneralSchedule(node x, slot t, int $nb_{channel}$, $List_C$)

$c \leftarrow 1$
test $\leftarrow$ false
repeat
  if No inter-cell interference no intra-cell interference c is available **then**
    schedule $\leftarrow$ c
    test $\leftarrow$ true
    /* the node x is scheduled for channel c and time t*/
  else
    if $c < nb_{channel}$ **then**
      $c \leftarrow c + 1$
    else
      test $\leftarrow$ false
    end if
  end if
until c=nbChannel —— test

6.9 Simulations and Results

In this section, simulations are performed using the MATLAB software based on many different scenarios. To this end, we validate the effectiveness of our trust method by comparing its performances with those of AND rule, OR rule and no-scheduling random approaches.

6.9.1 Spectrum Sensing (SS)

We will study the SS efficiency of the OR rule, AND rule models and our proposal on the correct decision, the error probability and the spectrum throughput in terms of the variation of faulty cognitive users sending non-accurate detections due to the presence of interference, multipath fading, hidden nodes, noises, shadowing, etc.

Foremost, we will define the two models OR rule and AND rule used in the cognitive environment to make the final decision about the channel availability.

- The OR rule: It decides that the channel is occupied if any of the users detects a signal.
- The AND rule: It decides that the channel is available if all users have detected a signal.

The three studied metrics for the spectrum sensing performance in the simulation experiments are:

1. The correct decision probability Pd: It is the probability that the DFC makes the correct decision about the PU activities.
2. The Error Probability P_{Err}: It is equal to the following equation:

$$P_{\text{Err}} = P(H0) * Pa + P(H1) * Pm \tag{6.11}$$

   where $P(H0)$ and $P(H1)$ are the probability for the PU to be absent/present, and Pa and Pm are, respectively, the probability of false alarm and missed detection.
3. The sensing throughput: It defines the successful transmission rate defined by the following formula:

$$\text{Th}_{\text{s}} = \frac{\text{size(message)} * \alpha}{\text{Period}} \tag{6.12}$$

where α is the number of the spectrum detection with success.

In this scenario, we will vary the number of faulty CR users in the network between 10 and 80%. Fig. 6.2 illustrates the variance of the correct decision probability in terms of the increased number of faulty users. We note that our proposal maintains the highest probability during the simulation which equals to 1 with 10% faulty users. Contrariwise, the correct decision probability progressively decreases with

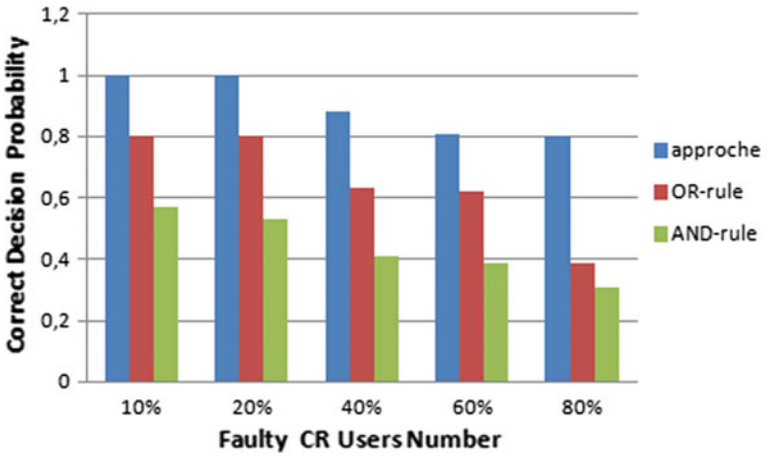


Fig. 6.2 Correct decision probability

the increasing number of inaccurate detection. Figure 6.3 portrays the influence of the erroneous detection number sensing by the faulty user on the error probability. Our proposal also outperforms the two models by maintaining a lower probability of error with the high number of faultiness which equals to 0.15 with 80% faulty users. Finally, Fig. 6.4 portrays the influence of the faulty user's number on the spectrum sensing throughput. We found that the higher number of faulty detections in the network cannot affect our proposal throughput. Indeed, our scheme can resist against a large number exceeding 80% of users. The throughput of the models without security falls heavily and cannot exceed 0.6.

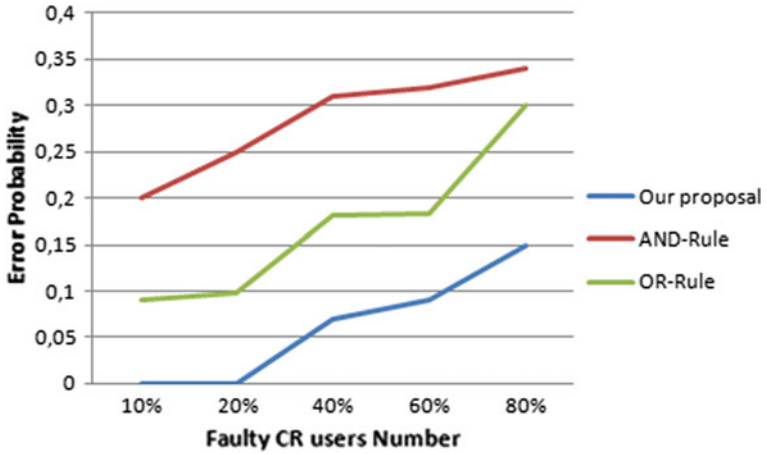


Fig. 6.3 Error probability

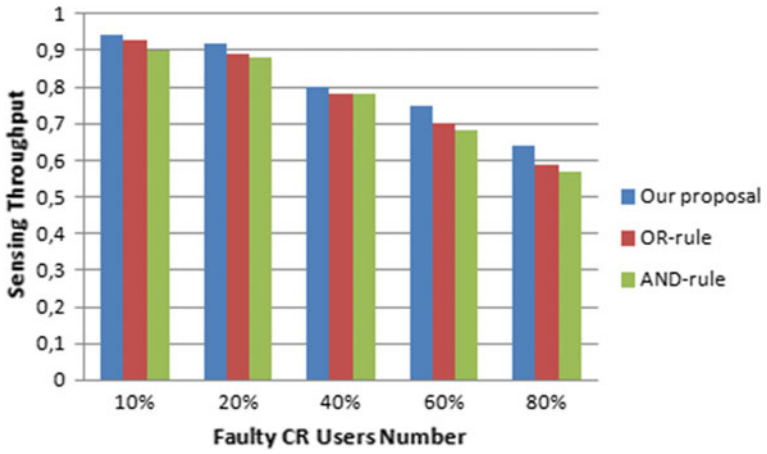


Fig. 6.4 Spectrum sensing throughput

As a conclusion, our approach outperforms the existing models. This is explained by the fact that our proposal takes into account the detection of faulty users sending inaccurate detections by using the trust scheme and the two game-based models. As a result, the network becomes more secure and more robust.

6.9.2 Data Routing

Thus, our proposal performances are studied and compared through three major metrics:

1. Mean throughput ratio is defined as the success transmission rate which is equal to the number of messages transmitted with success to the DFC.
2. Mean residual energy ratio is one of the must crucial metrics to validate the system efficiency computed as the remaining energy ratio of all nodes after the data transmission phase.
3. Average packet transmission delay: The delay for a packet transmission can be caused when the unoccupied channel is busy in serving other packets that arrive earlier or when there are no available channels during the reserved time interval. The delay is equal to the waiting time average of all nodes.

Firstly, we start by studying the impact of node number variation on the network performance. Figure 6.5 portrays the residual energy measured at the end of simulations which increases with the high number of nodes and reaches a very important value rate equal to 0.9 by our model. However, the residual energy evolution of the classical approach is not consistent. While the node number reaching 90%, the random approach cannot support the high number of traffic generated by the nodes and the remaining energy ratio falls drastically to 0.4. We find that our hierarchical

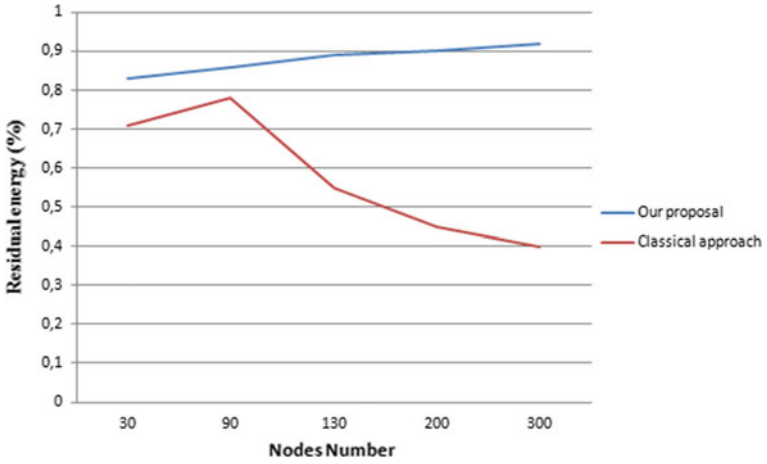


Fig. 6.5 First scenario: residual energy

proposal consumed less energy than the classical random approach. Figures 6.6 and 6.7 show the influence of node number variation on the delay and the throughput metrics. Due to the increasing high number of cognitive sensors competing to exploit the limited spectrum resource, the average delay value grows equivalently affecting the throughput of network basically resulting from the packets loss. In fact, in Fig. 6.6, our proposal outperforms the no-scheduling model by leading to a lower value of delay equal to 30 ms at the end of simulations which is explained by the fact that our opportunistic scheduling model aims at minimizing the waiting time of the active nodes having urgent data. Figure 6.7 portrays that the throughput and delay

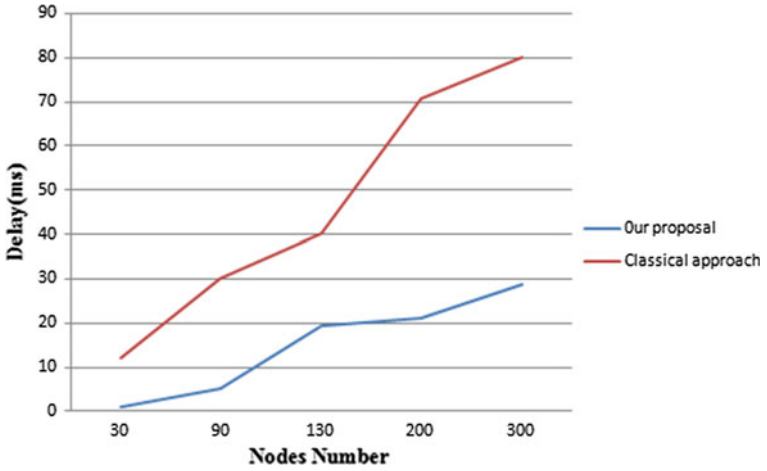


Fig. 6.6 First scenario: delay

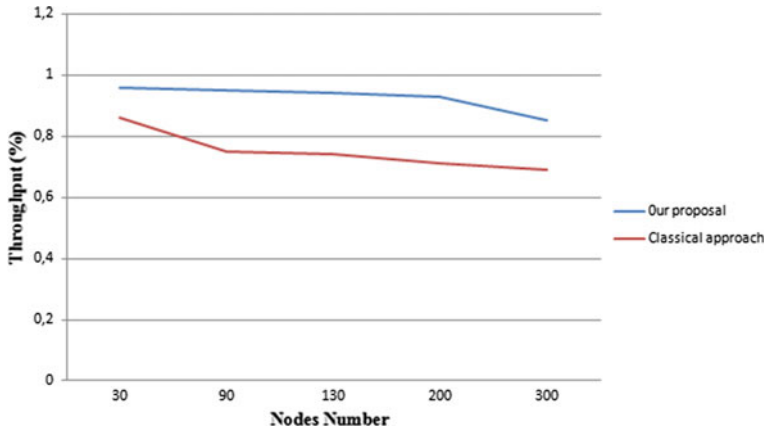


Fig. 6.7 First scenario: throughput

are inversely proportional. Thus, we find that our approach has the higher throughput level equal to 0.9 compared with the classical no-scheduling method. In conclusion, the proposed scheme outperforms the classical model in terms of the three metrics described above. In the next step, we estimate by simulation the optimal number of unoccupied channels to achieve the highest performance with the lowest resource allocation. Hence, we fix the number of SUs to 200 nodes and we vary the number of available channels in the network. Figures 6.8, 6.9 and 6.10 illustrate the variance of the residual energy, throughput and delay in terms of the increasing number of available channels in the network. Thus, our approach maintains the highest values during the simulation. With a high number of channels, the nodes can be served by transmitting their packets with a low delay (see Fig. 6.9). All figures show that the performance metrics discussed above become more and more stable when the

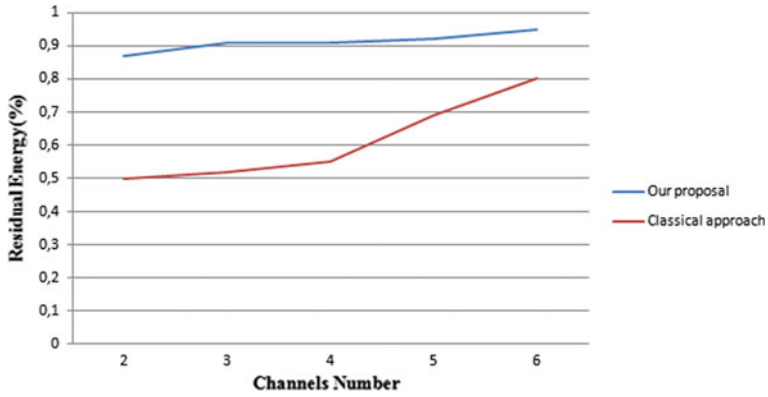


Fig. 6.8 Second scenario: residual energy

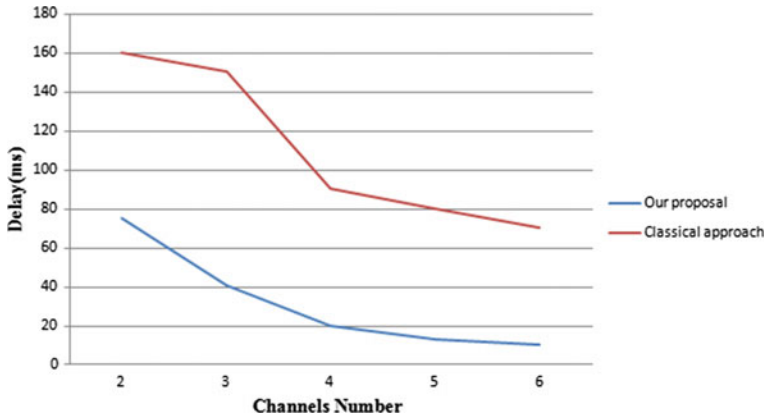


Fig. 6.9 Second scenario: delay

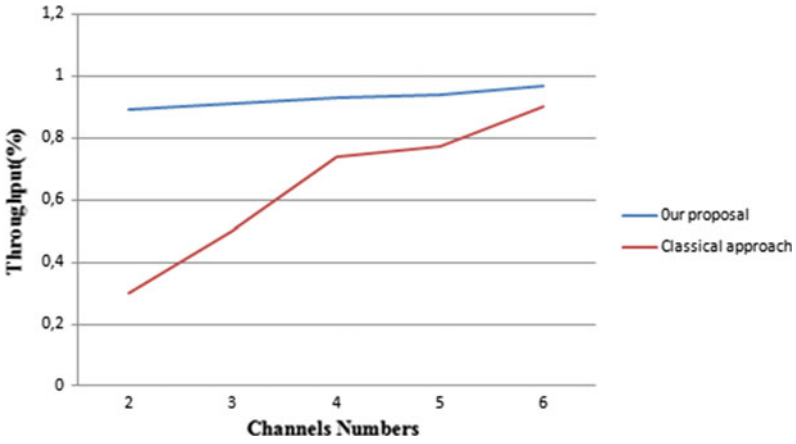


Fig. 6.10 Second scenario: throughput

number of channels exceeds 4. As a conclusion, the network can achieve efficient results in terms of the studied metrics with a number of channels equal to 4.

6.10 Conclusion

In this chapter, we presented a trustworthy novel and exhaustive design for the integration of cognitive radio in the IoMT application to mitigate the interference by providing robust cognitive communication. Firstly, we defined an efficient hierarchical model to facilitate the medical data transmission and to optimize the overall consumed energy. Moreover, we defined a trustworthy cooperative SS scheme based

on the trust system in order to enhance the spectral sensing accuracy leading to network performance improving. Then, we introduced our opportunistic channel assignment system based on competitive game models to optimize the resource allocation by defining two different priorities for the active and passive sensors. We compared, by simulations, our proposed hierarchical schema with the random distribution with classical scheduling system, OR rule and the AND rule models. We found out that our model improves the energy conservation, the data and sensing throughput and the correct detection probability. On the other hand, our proposed model enhances the network performance by reducing the transmission delay and the error probability.

References

1. Farahani B, Firouzi F, Chang V, Badaroglu M, Constant N, Mankodiya K (2018) Towards fog-driven IoT eHealth: promises and challenges of IoT in medicine and healthcare. FGCS 78(2)
2. Mitola J (2000) Cognitive radio: an integrated agent architecture for software defined radio. Royal Institute of Technology; [3] Dong Z, Sengupta S, Anand S, Hong K (2011) Cognitive radio mobile ad hoc networks in healthcare. Springer (2011)
3. Patil VM, Patil SR (2016) A survey on spectrum sensing algorithms for cognitive radio. HMI; [6] Pandit S, Singh G (2017) Spectrum sensing in cognitive radio networks: potential challenges and future perspective, spectrum sharing in cognitive radio networks
4. Kreutz D, Ramos F, Verissimo P, Rothenberg C, Azodolmolky S, Patil VM, Patil SR (2015) Software-defined networking: a comprehensive survey. Proceedings of the IEEE 103(1)
5. Wyglinski AW, Orofino DP, Ettus MN, Rondeau TW (2016) Revolutionizing software defined radio: case studies in hardware, software, and education sign in or purchase. CM 54(1)
6. Ali Khan A, Husain Rehmani M, Rachedi A (2017) Cognitive radio-based internet of things: applications, architectures, spectrum related functionalities, and future research directions. IEEE Wireless Commun 24(3)
7. Khan AA, Rehmani MH, Rachedi A (2016) When cognitive radio meets the internet of things? IWCMC
8. Hu X, Cheng J, Zhou M, Hu B, Jiang X, Gua Y, Bai K, Wang F (2018) Emotion-aware cognitive system in multi-channel cognitive radio ad hoc networks. IEEE CM 56(4)
9. Nanthini SB, Hemalatha M, Manivannan D, Devasena L (2014) Attacks in cognitive radio networks (CRN)–A survey. Indian J Sci Technol 7(4)
10. Marinho J, Granjal J, Monteiro E (2015) A survey on security attacks and countermeasures with primary user detection in cognitive radio networks. EURASIP (2015)
11. Bennaceur J, Idoudi H, Azouz Saidane L (2018) Trust management in cognitive radio networks: a survey. Int J Network Mgmt
12. Dong Z, Sengupta S. Anand S, Hong K (2011) Cognitive radio mobile ad hoc networks in healthcare. Springer
13. Phunchongharn P, Hossain E, Niyato A (2010) A cognitive radio system for e-health applications in a hospital environment. Wireless Commun
14. Al Mamoon I, Muzahidul Islam AKM, Baharun S, Komaki S (2016) A novel medical priority aware transmission mechanism for cognitive radio based hospital. TENCON
15. Ouattara D, Minh Thao Q, Krief M, Aymen Chalouf M, Khalif H (2013) Mitigating the hospital area communication's interference using cognitive radio networks. In: International conference on e-health networking, applications and services

16. Van Moer W, Bjorsell N (2012) Saving lives by integrating cognitive radios into ambulances. MEMEA
17. ul Hasan N, Ejaz W, Atiq MK, Seok Kim H (2017) Energy efficient error coding and transmission for cognitive wireless body area network. Energy Efficient Netw 30(7)
18. Jaafar Ali M, Moungla H, Younis M, Mehaoua A (2017) IoT enabled channel selection approach for WBANs. IWCMC
19. Bagaa M, Chelli A, Djenouri D, Taleb T, Balasingham I, Kansanen K (2017) Optimal placement of relay nodes over limited positions in wireless sensor networks. Trans Wireless Commun
20. Mabrouk O, Idoudi H, Amdouni I, Soua R, Minet P, Saidane L (2014) OTICOR: opportunistic time slot assignment in CRSNs. AINA
21. Clark MA, Psounis K (2017) Equal interference power allocation for efficient shared spectrum resource scheduling. Trans Wireless 16(1)

Chapter 7
E-Health Threat Intelligence Within Cyber-Defence Framework for E-Health Organizations

Arif Sari and Joshua Sopuru

Abstract In recent years, scholarly work on cybersecurity in smart health has gained substantial attention from both practitioners and scholars. This is primarily due to the rapid growth in the field of information, communications and technology, protocols, an important aspect of smart health communication infrastructure. The smart health communication infrastructure is solely developed to provide data communication for specific networks such as wireless body area network (WBAN) which is developed for the health sector. The modern healthcare service delivery eliminates the need for real-time inspection of elderly and attention-need patients; that is, medical experts can monitor such people from a remote location through e-health communication infrastructure. The developed communication infrastructure is used by e-health organizations to store, process or transfer patient's data which has high priority and requires confidentiality. The infrastructure used by e-health organizations must restrict unauthorized access to patient data against any intruder. e-health organizations are a major target for hackers as they hold a huge amount of private data as a source of wealth of information. The proposed security solutions for e-health organizations require specific policy developments and propose solutions for specific security layers. The smart, scalable and adaptable solutions are proposed by researchers to overcome several security challenges in e-health organizations. Some of the proposed solutions provide open use and sharing of critical e-health data without compromising patients' rights to privacy and confidentiality. The deployment of these solutions faces several problems since hackers targeting network layer of these models. Development of new attack methodologies and technological enhancements strengthens hackers to attack with different motivations and compromise e-health organizations' private data. For this reason, a new security framework is necessary for e-health organizations' communication infrastructure. The privacy of the patient's health data must be carefully addressed while developing a new framework. In order to maximize the healthcare quality and minimize the e-health cost, the ultimate goal of this chapter is

A. Sari (✉) · J. Sopuru
Department of Management Information Systems, School of Applied Sciences, Girne American University, Kyrenia, Cyprus
e-mail: arifsarii@gmail.com

J. Sopuru
e-mail: sopuru.senya.joshua@gmail.com

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_7

to expose the limitations in the current e-health organization cybersecurity solutions and provide a new security framework to highlight existing gaps in communication infrastructure of e-health organizations. The comparison of cryptographic attacks against encryption algorithms to secure communication infrastructure, latest zero-day attacks in e-health sector, network layer attacks to e-health organizations and e-health threat intelligence will be investigated within the scope of this chapter. The e-health threat intelligence will be the main contribution of this chapter since threat intelligence provides insight about the possible threat and ensures that e-health organization can defend against zero-day vulnerabilities and protect the patient and other staff personal identification information.

7.1 Smart E-Health Systems

Wireless sensor network (WSN) is needed for the implementation of a secured and smart e-health system. These systems can be extended to integrate sensors for the implementation of the Internet of things, which is the driving technology empowering e-health. IoT which comprises embedding computing devices on objects for the sole purpose of communication (data exchange) via the Internet has given birth to a wide range of possibilities in the health sector. Devices armed with the responsibility of monitoring patients' body can collect and transmit vital signals via a secured network. According to Bhat et al. [1], IoT offers great prospects to e-health as embedded devices are used in health systems to monitor various reactions in a patient's body.

Ubiquitous sensor network (USN) and wireless sensors can provide efficient health care by ensuring constant monitoring of patients, emergency response, clinical diagnosis, etc. [2]. Treatment of patients will shift from hospital-based to home-based as the underlying technologies empowering e-health continue to grow. According to Rahmani et al. [3], it is predicted that before 2030 this will be achieved. Because of this, the convergence of the Internet of things, smart devices, cyber-defence systems and e-health threat intelligence should be deeply considered.

Gateways which serve as interfaces between smart homes/hospitals and personal/local area networks must also be carefully considered during implementation to ensure optimum security of individual data. Due to the high processing power of these gateways, transfer of data between hospitals and patients is done in real time. This advantage, however, can be capitalized upon by fraudulent players to cause harm to the overall network. The availability of high processing power of gateways can also be further capitalized upon by adding some levels of intelligence, thereby widening its basic functions to include that of an intelligent server. Rahmani et al. [3] also assert that gateway capabilities can be used to enhance IoT architectural design to improve reliability, performance, interoperability and security of healthcare networks.

7.1.1 E-Health System Architecture

Rahmani et al. [3] in their work presented a basic architecture for an e-health system. In their design, information related to patient's health is collected by a sensor implanted into the body, and complementary data such as time, date, patient location and temperature can be easily recorded and transmitted in real time to a monitoring system. Unusual patterns are identified through context awareness, these patterns are collected, and inferences are carried out to determine the patients' situation. Their design includes the following axiom components:

1. Medical Sensor Network: A network of sensing and identification devices with communication ability; some of these devices are implemented into the patient and some embedded on medical equipment and prescription. The devices collect context and biomedical signals used for either treatment or diagnostic purposes [4]. Collected signals are then transferred wirelessly or via a cable connection to the gateway. According to Schmidt et al. [5], attaching wearable wireless devices that collect vital health signals on patients would enable nurses, doctors and all caregivers monitor continuously the state of a patient. Example of such sensing devices includes open artificial pancreas system (OpenAPS), continuous glucose monitor (CGM), connected inhalers, etc. Sensors embedded on medicines help doctors determine whether patients adhere to prescription or not.
2. Smart e-health gateway: Gateway serves as a connection point between the sensor network and the Internet. Equipped with the ability to accommodate different connection protocols, data can be received from different sensing networks. One of the major functions of the e-health gateway is to collect aggregated data from different sources, filter the data and prepare it for transmission by dimensionality reduction (the breaking down of data into easily classified and usable form).
3. Back-end system: Two components make up the back end of the system: a cloud computing area and a local switch. The cloud computing platform is made up of data warehouse, broadcasting functions and big data analytics servers. A GUI web client is attached to the system for visualization purposes.

The authors assert that the gateway of an e-health system plays a vital role in determining several aspects of the system. Some of these aspects, one of which is the focus of this chapter, are the e-health threat intelligence.

Computer networks need a secured defence mechanism that can effectively respond to threats and defend system infrastructure [6]. The most important consideration in implementing a health system is security. Sensor nodes in most cases disseminate data in plain text format which can easily be intercepted by unauthorized players. In some cases, however, data are disseminated in encrypted form making it difficult for unauthorized access. In addition, most gateways are built on the Linux kernel which comes by default with the iptables configuration.

The iptables is a Linux security feature where only a few ports are open (usable), thereby restricting access to other ports. However, despite the advances in ensuring a secured e-health system, Wu et al. [7] emphasize the need for more work in achieving a secured e-health system.

7.1.2 The Current State of Cybersecurity in E-Health

According to Wu et al. [7], major cyber-attacks on health organizations grew in 2016 with 63%. The US Department of Health and Human Services reported a steady growth in the number of attacks with malware shutting down operations of some hospitals.

DDoS attack targeted the central health information systems of city hospital in Finland which resulted to a total shutdown of hospital operations for two days. Health information systems of Lukas Hospital in Germany was hacked and shutdown by ransomware that interrupted the patient appointments, diagnosis and treatment services [7].

Government organization such as the European Union (EU) has made reasonable contributions to ensure improved information security. In September 2005, the European Network and Information Security Agency (ENISA) began operations with the objective of optimizing the security of information systems within EU member states.

As cited by Wu et al. [7], some of the major contributions of ENISA towards cybersecurity in health care include ensuring security/resilience for e-health services/infrastructures and smart hospitals with cybersecurity. Some ongoing projects are NISD implementation and cloud security (OES-DSP dependency).

Despite the commitments of ENISA [8], serious challenges still abound. According to the situational report of the ENISA (published in 2017), because most hospitals do not have needed infrastructure, cybersecurity maturity has been low. Also, the poor security awareness of stakeholders is an impediment to attaining cybersecurity in the health sector [8].

The graph below presents the most important security challenges in e-health infrastructures and systems (Fig. 7.1).

Assenting to the ENISA report, Faruki et al. [9] in their paper reveal among several other reasons that poor defences and rich data banks of health care are the two rudimentary reasons why health care is a favourable playground for a cyber-attack. According to them, despite how critical cybersecurity is to patients and its

Fig. 7.1 Cybersecurity challenges [8]

overall importance in health care, its benefits to e-health have since history been insufficiently mined.

7.2 Review of E-Health Cybersecurity Models

With the spread of smart devices, there has been a change in the way healthcare organizations collect data. Wearable devices with sensors provide a large volume of data that need to be efficiently collected, stored and information inferred from it whenever needed [10]. As this large volume of data is generated, it is expected that only authorized individuals have access to it as it contains sensitive and confidential information. However, this is not usually the case as many unauthorized individuals are interested in such confidential information as contained in e-health data. Several researchers have investigated different cybersecurity models implemented on e-health systems.

Encryption has been the fundamental solution used for e-health systems as it has proven efficient in simple e-health systems. When applied to more complex systems, however, it is not as straightforward as compared to simpler systems (Burguera et al. [11]

A modular access control (MAC) was developed by Morchon and Wehrle for healthcare applications. In their model, the RBAC model was extended because of the distribution of access to nodes and storing medical context to influence a patient's situation. Their design achieved safer deployment of data from medical sensor networks.

Ramlil et al. proposed a security framework that is based on the authentication of data through biometrics within a wireless body area network (WBAN). The main idea is that the senders' electrocardiogram (ECG) factor can be used as a key to making sure that the mixture of data that belongs to the patient can be avoided. Hence, it is based on biometrics, and a different patient has his/her biometric attributes. This will improve the functionality and distribution of cryptographic key and reduce mistakes made computationally. The framework saves the usage of lots of resources and employs effective security measures. The setback of this framework narrows to the fact that sensors are used for authentication which tends not to be a significant help to the limited resources.

Amini et al. proposed to design an effective lightweight security model. The model evolves around different examined security protocols which include MiniSec, RC4-based, TinySec and LLSP which are tagged along certain cipher algorithms (AES, RC4 and Skipjack). Combination of different types of attacks (spoofing of sensors, data loss, replay and eavesdropping) was done by the authors. More so, they applied the cipher algorithm and discovered that in order to fulfil confidentiality based on ROM, RAM and clock cycles per byte (CPB), cypher algorithms like Skipjack and RC4 are most suitable. Though they came out with different advantages, they failed to dive deeper into other threats that oppose the security.

Gajanayake et al. presented a model of privacy-oriented access control which is aimed at satisfying e-health's demand. In this model, the combination of three different access models (MAC, DAC and RBAC) into a module projects the ability of healthcare professionals and patients to ascertain and set access privileges. On different occasions, the module has been examined to define various accesses to data and policy settings. This model demonstrates that it can be a stand-alone security model.

Maw et al. designed an adaptive access control model which can give an access with rights to medical data in WSNs and BSNs. The designed model put into consideration behaviour and privilege overriding, so clients have access to interrupt or stop a denial of access whenever there is an unexpected occurrence. In this model, users do not have to get privilege or authorizations; hence, their sessions are based on behaviour trust in respect of users, time action and location. In contrast, the limitation is that the detection mechanism to check the user's data access when an unexpected event occurs is not implemented.

7.2.1 E-Healthcare Encryption Algorithms

In e-health, data can be encrypted such that only authorized individuals who have the secret key can decrypt such data. Encryption of data has been used over the years in medical fields to restrict access to patients' data to only authorized doctors. The AES encryption algorithm is a standard widely used in encrypting and decrypting important information. It uses an algorithm called block cipher to generate secure data. Figure 7.2 shows a basic structure of how AES works.

As shown in the figure above, patients' data are encrypted using a secret key. The cipher converts these pairs into a ciphertext which becomes impossible to decrypt without the key.

A modified AES algorithm was designed by Shabtai et al. [12]. In their algorithm, a square matrix (Polybius) was used with a 4×4 matrix key. The first step of encryption

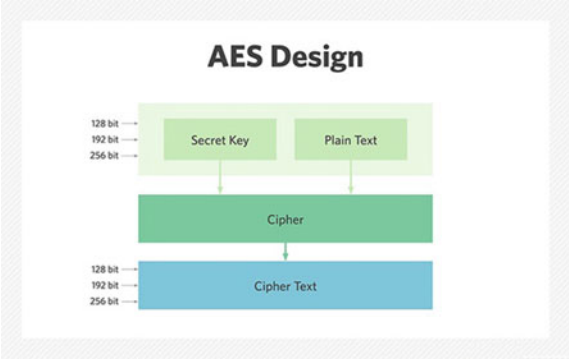


Fig. 7.2 Standard AES design

involves getting the XOR of the input based on a 4 × 4 key matrix. The input is then transformed to a nonlinear form. Transformation is done by generating the inverse. Diffusions are then created by shifting the rows of the 4 × 4 matrix after which columns are mixed following the Galois field (finite number of elements contained in a field). The advantage of this improved AES is that at the completion of the padding, cracking of the encryption becomes impossible.

Hoang et al. improved the efficiency of input encryption at a medium access control (MAC) level. Using a 128-bit key, they successfully implemented the AES-CCM IP which consumed lower power and was highly efficient.

One common limitation of these algorithms is the increase in processing time. In Puneet Kumar and Shashi B. Rana model, the use of a nonlinear S-box will increase memory requirement as rounds of the OTP increase [5].

The fingerprint-based mutual authentication is another protocol proposed. Designed based on the EPC-CIG2 framework, the model conforms to the EPC standard. One major disadvantage of this model is that eavesdropping has not been effectively curbed [7].

7.2.2 Types of Attack

Since most sensing devices are defined by rigid constraints in terms of the computing resource, memory size and energy, many of them cannot accommodate complex security algorithms; thus, designing effective security algorithm for e-health sensory network is crucial. It is important we understand the architecture of IoT devices. IoT solutions are established on an architecture with three basic layers which are perception layer, application layer and network layers.

Sensors and actuators make up the perception layer. This layer handles things, for sensing devices. The application layer is the brain of the IoT. On this layer, logic and applications are defined. As the perception layer collects data, the application layer performs logical operations on the data. The network layer, on the other hand, is responsible for the transmission of information gotten from the perception layer. Security can be bridged, and data compromised within a sensing network if proper security considerations are not put in place. We will classify security attacks into five types: physical, side channel, cryptanalysis, software and network.

Physical attacks: These attacks are difficult to perform as they attack physical hardware components.

Side channel attacks: They are usually encryption-based attacks, the base in the retrieval of information from an encryption device. These types of attacks are made possible when there is a lag in encryption.

Cryptanalysis attacks: These attacks focus on trying to break encryption keys used in encryption. Software attacks are currently the major type of attacks on e-health

Table 7.1 Examples of active and passive attacks

Active attacks	Passive attacks
Masquerade	The release of message content
Modification of messages	Traffic analysis
Repudiation	
Replay	
Denial of service	

systems, and this form of attack involves exploiting software programs, using malicious programs to gain access for destructive purposes.

Network attack: It has to do with communication system vulnerability, and it occurs when an attacker capitalizes on the vulnerability of a network. Example of such attack includes denial of service (DOS) attack.

For a better understanding of how these attacks work and how they affect e-health, we will consider cybersecurity attacks as active and passive attacks.

Active attacks are attacks that seek to change system files or resources, thereby altering system operations. It involves changing or modifying streams of data involved in system operation. On the other hand, passive attacks are attacks which wish to learn about a system. They do not modify system resources but only gather and make use of critical information about the system (Table 7.1).

Masquerade:

Just as the name implies, masquerades take up a false identity. In a masquerade attack, an entity possesses as a different identity. This attack can be considered as personification (when a person pretends to be another person). For example, Bob pretends to be Lily in other to gain access to the communication between Lily and John. Before Lily and John begin communication, Lily must have fulfilled some security requirements as already defined by the system. Bob attempts to clone (or duplicate) Lily's authorization requirements and gains her privileges to be able to talk with John. John on his own end believes he is communicating with Lily as Bob has fulfilled all security requirements (Fig. 7.3).

Modification of messages:

In the Byzantine problem, a major issue addressed is the integrity of information received amid attackers. The Internet is filled with legitimate and illegitimate (attackers) users. Sending a vital message over an unsecured network can be very disastrous. The e-health network is a crucial network as messages involving life and death can be transmitted over the network. Modification of messages is a form of attack where an attacker intercepts a message, modifies the message and forwards the modified message to the originally intended recipient. For example, Doctor John prescribes medication for his patient Lily, "give Lily two doses of penicillin". Before this message gets to Lily's home nurse, an attack Bob intercepts the message modifies it to read "give Lily four doses of penicillin" and forwards it to Lily's nurse. On

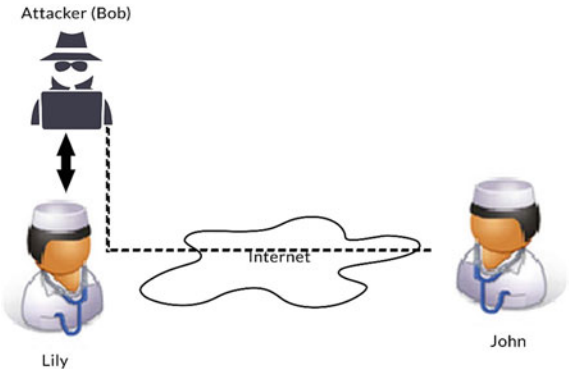


Fig. 7.3 Masquerade attack

the other hand, modification of message attack can come in the form of an attacker modifying the packet header of a message and redirecting the message to a different destination other than the original destination (Fig. 7.4).

Repudiation:

A repudiation attack can be done by the sender or the receiver. The attacker, either the sender or the receiver, denies sending or receiving a message. A secured system tracks and monitors user activities. This monitoring is usually managed in the form of session. A repudiation attack is an attack where a system denies authoring an event; as a result of this denial, the system does not keep proper track of controls or user events, giving attackers opportunity to jeopardize system security.

Replay:

It is the temporal capture of a transmitted message by an unauthorized user. The message is delayed temporally by the attacker and retransmitted to a single or multiple user. Unless handled properly, networks and users will see a message sent from a replay attack as legitimate. An example of a reply attack is when a legitimate message is sent and acted upon by the recipient and after some time the same message is resent

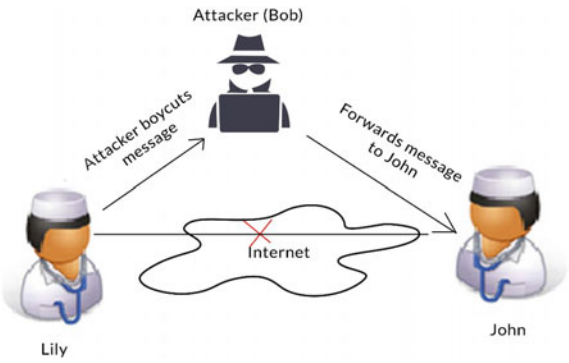


Fig. 7.4 Modification of message attack

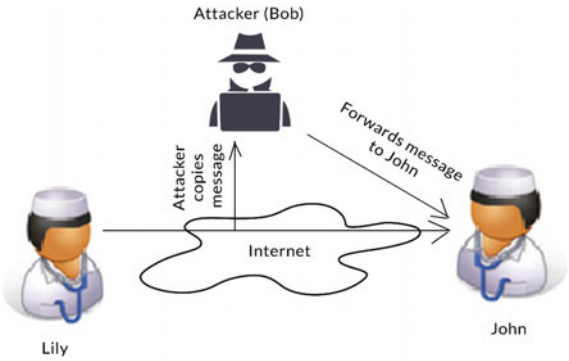


Fig. 7.5 Replay attack

by an attacker to the same user for fraudulent reasons. The use of timestamps and strong digital signatures can help prevent replay attacks (Fig. 7.5).

Denial of Service:

Denial of service also called DOS attack is when an attacker denies legitimate users' access to specific system resources or devices. This attack involves flooding of a server with spam request that generates traffic the server cannot handle. The victim's system utilization exceeds its capacity and therefore becomes unable to service legitimate users. DOS attack is more severe than many other attacks resulting in a server crash because, while other attacks can be handled with a server reboot, a DOS attack might not be resolved by just a reboot. Recovery becomes even more difficult if the attack comes from distributed sources; in this case, the attack is referred to as a distributed denial of service (DDOS) attack. DOS and DDOS attack can lead to a catastrophe in the e-health system as most services on the network are time sensitive.

Network configuration vulnerability in traffic handling is the main breeding ground for DOS and DDOS attack. For example, if the traffic handling protocol of a network is blind towards multiple transfers of large packets of data, an attacker can capitalize on that vulnerability by bombarding the server with multiple large packets of data. One of the practices of Internet Service Provider (ISP) to mitigate a DOS or DDOS attack is the use of load balancers. Load balancers distribute loads equally among many servers, thereby reducing stress on one server (Fig. 7.6).

The release of message content:

Information transferred via a network may contain confidential messages intended solely for the intended recipient. However, a passive attacker can monitor the contents of messages transmitted over a porous network. The sender and the receiver are usually unaware of the privacy bridge, and as a result, it can lead to more divulge of private information.

The diagram below shows the release of message content attack (Fig. 7.7).

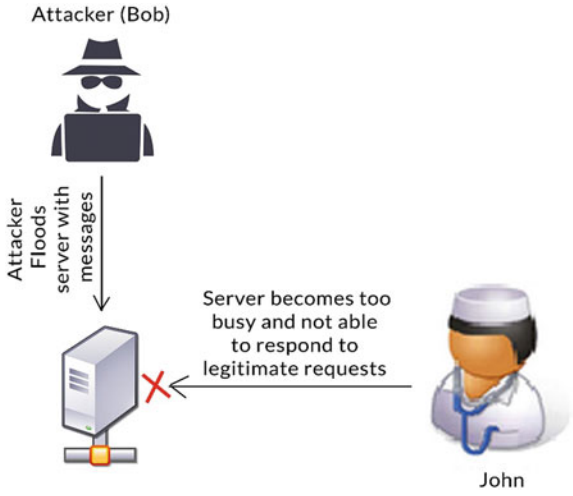


Fig. 7.6 Denial of service attack

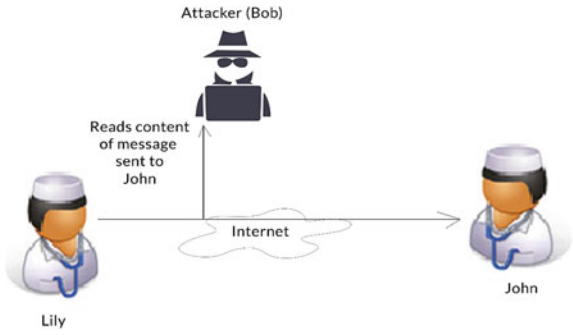


Fig. 7.7 Release of message content

Traffic analysis:

Traffic analysis as the name implies is when an attacker spends the time to analyse the traffic of a network for illegal purposes. Packet size, senders' IP, receivers; IP, port numbers, duration of the connection, etc., can be effectively monitored by an attacker. Supposing messages are well encrypted such that even if an attacker intercepts it, he cannot break the encryption algorithm, the attackers may decide to monitor the traffic so as to gain more knowledge on the pattern of information and how to penetrate the conversation.

This type of attack involves intelligence. Attackers engage different means in other to derive intelligence from the data collected during the analysis. Metadata collected include endpoint addresses, timing (duration and sequencing), location, etc.

7.2.3 E-Health Threat Intelligence

There is a need for every organization to keep relevant security information. Such security information is needed to protect the organization form internal or external threats. Information kept may include tools, policies or processes developed to collect and analyse data.

According to Shabtai et al. [12], security intelligence is the collection, normalization and analysis of user-generated data, infrastructure and application in real time that affects IT security of an enterprise. In e-health, threat intelligence is important. Bhat et al. [1] assets that gateways can be expanded to accommodate different intelligences, since gateways can accommodate complex logic which provides intelligence. In this chapter, we propose a framework that incorporates cyber-defence intelligence into the e-health network gateway.

Security intelligence is defined by some key principles.

Real-Time Analysis: Knowledge of happenings in the network is important in identifying threats. Log records keep information on happenings on the network, but a real-time analysis of current happenings in the network will help trigger important security information.

Pre-exploit Analysis: Pre-exploit vulnerability is blended with real-time analysis by modern security intelligence. This involves systems that identify risks early enough before they result in breaches. Loss because of an attack on health facilities can be avoided with such systems in place.

Actionable Insight: The saying goes "Prevention is better than cure". Identification of threats, removal of false positives and forwarding potential threats to analysts must be a proactive measure taken by a security intelligence solution.

Scalable: e-health devices produce a large quantity of data. Intelligent solutions are designed to handle these large data. Customized databases are implemented to collect and analyse in real time a large amount of data.

Data Risk and Security: Protecting the data of e-health applications is the main goal of a security intelligence solution.

Based on these key principles, we derive data that can be gathered and analysed for security intelligence in an e-health system. Further categorization based on i™ external data is shown in the Table 7.2.

Firewall logs: Firewall logging is important to track malicious activities within a network. As the first line of defence, the firewall stands as a protective shield against external threats. Blocking suspected traffic from accessing infrastructure, firewall ensures the security of the network. Deriving knowledge based on firewall logs, intelligence can be created from the traffic sources that were allowed and denied.

Firewalls work by inspecting source address, the intended destination/address and the targeted port number. Information on these addresses and the ports is considered important characteristic (tuple) to be evaluated based on a predefined set of rules. If the tuple bears information matching an approved connection, the connection is

Table 7.2 Internal and external data sources for e-health threat intelligence

Internal data	External data
Firewall logs	Top tier phishing indicators
Proxy logs	Brand abuse indicators
IDS/IPS logs	Malware campaign indicator
Endpoint security	Fraud payment logs
Employee directory	Actor intel
Network security logs	Staff asset
Fraud payment logs	IP reputation
Application inventory	Malware hash MD5
DNS/DHCP logs	

then granted to the connecting device. The rules of the configuration of a firewall are very important as it defines the intelligence on such a firewall. As stated earlier, the ability of the firewall to keep the record of not just allowed tuple provides wide resources for deriving intelligence.

Figure 7.8 shows an example of a firewall log and the data it contains.

This log keeps helpful information of date, time, action, protocol, source IP, destination IP, size of the packet, TCP flags, etc.

Analytics involves vulnerability check, management, identity access, data security and cloud computing.

Proxy logs: Proxy logs keep important information about access to the Internet. These logs contain information such as time, client, log-tag, size, request, URL, user id, and hierarchy. Self-organizing maps (SOMs) have been implemented in analysing web proxy data. According to Fei et al. [13], SOMs are important in network forensics

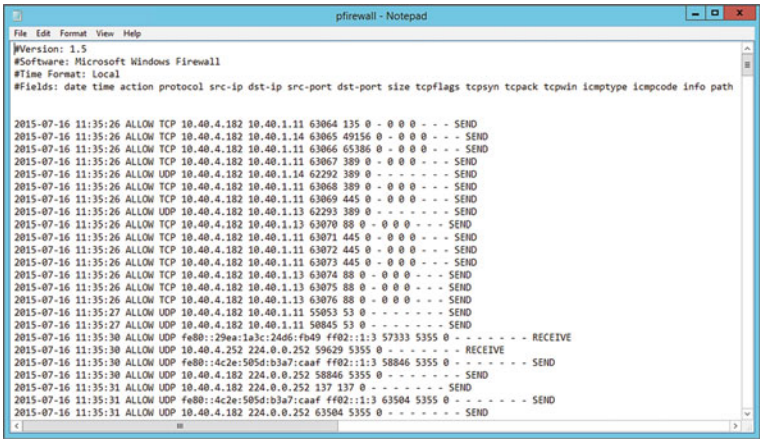


Fig. 7.8 Firewall log and data

in detecting unusual behaviour. Through the investigation of access patterns to a network, anomalous behaviour can be detected.

The data sources listed in the table above all play an important role in contributing to an intelligent e-health network. Deriving intelligence from these data will involve a correlation of events and some form of analysis.

An intelligent system must be able to understand events and effectively categorize them as authorized or unauthorized. For this, the threat intelligent system must derive knowledge from input data, analyse input data based on the stated algorithm, refer to a database of known events and causes of such events, and finally update its registry on the new event encountered. With a shared central cloud database of known attacks and description (in terms of data and parameters) of events, individual e-health centres can derive knowledge from the experience of other e-health centres. For example, the spread of the ransomware (Wanna Decryptor) attack to over 16 hospitals in the UK would have been curtailed to one hospital if such intelligent system was put in place.

Figure 7.9 shows a block description of multiple e-health centres extracting knowledge based on the experience of other centres.

It might be a daunting task for individual centres to detect and deal with all unknown attacks before it affects their e-health infrastructure, but the spread to other e-health centres can be curtailed for all unknown malware.

Filtering input data:

One of the major strengths of developing threat intelligence is the fact that no attacker can attack a centre without leaving behind trails of data. The strength of an algorithm to decipher hidden/suspicious patterns contained in a data stream will determine the level of intelligence attainable. In filtering input data, the algorithm must be able

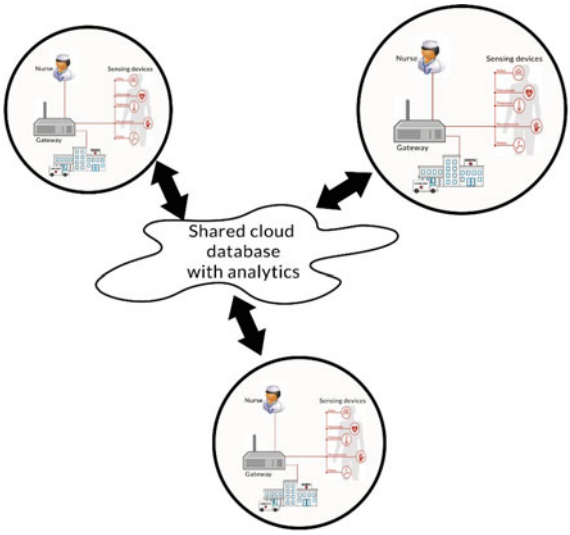


Fig. 7.9 Multiple e-health centres extracting knowledge

to categorize each connection information by making multiple inferences based on known connection parameters and related events triggered [14].

For example, a packet containing the following network parameter was sent to an e-health centre. Using different network connection query commands, we can retrieve information about the connection. For example, the ipconfig, netstat, host, nslookup, dig, nmap, etc., are network connection query commands we can use to gain a better understanding of the connection.

Information we can get from these commands is shown below:

Ipconfig: The ipconfig displays configuration data with packet counts telling you how busy a network interface is

$ ifconfig eth0

eth0 Link encap:Ethernet HWaddr 00:1e:4f:c8:43:fc
inetaddr:192.168.0.6
Bcast:192.168.0.255
Mask:255.255.255.0
inet6 addr: fe80::b44b:bdb6:2527:6ae9/64 Scope:Link

Netstat: Network connections and routing information are provided by the netstat command

$ netstat -rn

Kernel IP routing table

Iface					
0.0.0.0	192.168.0.1	0.0.0.0	UG	0 0	0 eth0
169.254.0.0	0.0.0.0	255.255.0.0	U	0 0	0 eth0
192.168.0.0	0.0.0.0	255.255.255.0	U	0 0	

Host: Like the *nslookup* command, the host looks up a user's IP address that is remotely connected to a network

$ host *world.std.com*

world.std.com has address 192.74.137.5
world.std.com mail is handled by 10 smtp.theworld.com.

Dig: The dig command gives detailed information about the remote connection. It tells how long a query takes and provides details about the name server.

$ dig *world.std.com*

≪≫ DiG 9.10.3-P4-Ubuntu ≪≫ world.std.com
-≫HEADER ≪- opcode: QUERY, status: NOERROR, id: 28,679
flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 1
OPT PSEUDOSECTION
EDNS: version: 0, flags:; udp: 512
QUESTION SECTION
world.std.com. IN A
ANSWER SECTION
world.std.com. 78146 IN A 192.74.137.5
Query time: 37 ms
SERVER: 127.0.1.1#53(127.0.1.1)
WHEN: Mon Oct 09 13:26:46 EDT 2017
MSG SIZE rcvd: 58

Data retrieved from a connection must be analysed for effective categorization. Keeping records of all connection data and the events, such connections triggered on the system are important as it will help in the analyses and categorization of new connections.

The diagram below illustrates the connection log of an intelligent system. Each connection is mapped to a series of events triggered on the system with timestamps.

The system keeps a history of all events relating to a connection in the form of an infographic. Each event is given an identifier and can be queried in other to create knowledge about the established connection (Fig. 7.10).

This form of mapping allows individual parameter in the network connection information to be evaluated independently or in relation to other parameters. Also, as connection variables change, we can easily identify and understand changes in applications accessed and ports opened or closed. Proper record and presentation of

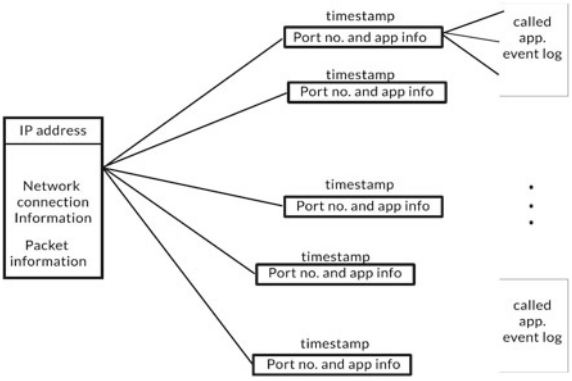


Fig. 7.10 Mapped network connection parameters with app events and ports

connection history and how it relates to network parameters will aid in detecting an attacker's connection.

7.2.4 Event Correlation

To have a safer and easier event correlation, data generated from multiple sensors are fed into a central analysis console (CAC). A detailed image of the event chain is generated as activities from different logs are collected.

7.2.5 Smart E-Health Gateway

The smart e-health gateway can be used to host defined security rules and protocols that can be used to ensure a secured e-health network [15, 16]. As different nodes connect to the gateway, the CAC collects and correlates data for making intelligent security decisions. The CAC is equipped with a combination of network security monitoring tool such as Sgull. The major duty of Sgull is to provide access to session data, real-time events, packet capture, etc. These signals are investigated for vulnerability, and when an input needs further investigation, Sgull makes available in a seamless manner data needed for decision-making.

The expansion of the smart e-health gateway to incorporate CAC will not only make readily available data needed for intelligence but also aid in decision-making as to what to do when new threat variants are discovered. The diagram depicted in Fig. 7.11 is a block representation of threat intelligence within a cyber-defence network for e-health organizations [17–19].

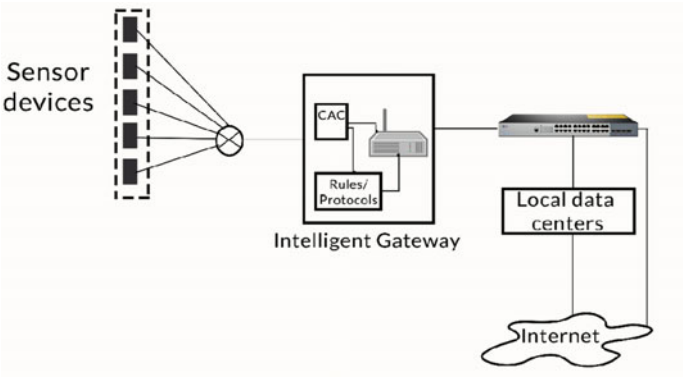


Fig. 7.11 Smart gateway with threat intelligence

7.3 Conclusion

Smart health communication infrastructure is solely developed to provide data communication for specific networks such as wireless body area network (WBAN) which is developed for the health sector. Sensing devices that collect real-time data are prone to security attacks, and such can jeopardize the overall e-health system. Different security measures such as message encryption have been implemented to secure data exchange on an e-health network. Despite these efforts, the security of e-health facilities is still a challenge as attackers' have constantly divers' alternative methods to jeopardize network security. In order to effectively harness the capabilities of an e-health network and maintain optimal security, intelligent modules can be embedded on gateways with the aim of achieving threat intelligence, thereby ensuring a more secured network. Just like every other intelligent framework, an intelligent gateway utilizes data generated from its network and those gotten from other networks to make intelligent decisions. Suspicious behaviour observed from devices connected to the gateway can be effectively monitored and intelligent decisions taken. Since the gateway serves as access point between sensing devices and healthcare facilities, implementing decentralized gateways capable of communicating with other gateways will help evade total shutdown as a result of an attack.

Utilizing data generated from different sensing devices connected to a gateway, we compared them to data gotten from same devices but connected to other gateways. Suspicious tuple is isolated by the CAC and data generated seamlessly made available for further analysis. If a tuple is identified as risky, the CAC triggers a response call denying access to e-health facilities until such threat is handled. With this design, intelligent rules/protocols can be defined to serve as proactive measures against unauthorized access to health facilities.

The ability of authorized persons with administrative privileges to define security rules/protocols thereby providing opportunity for personalization is a recognized advantage of an intelligent gateway. Another advantage is its collaboration with other gateways to quickly identify and eradicate new threat entrants before spreading to multiple e-health facilities.

References

1. Bhat MI, Ahmad S, Amin A, Ashraf S (2017) e-Health with internet of things 6(6):357–362
2. Pantelopoulos A, Bourbakis NG (2010) A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans Syst Man Cybern Part C (Appl Rev) 40(1):1–12
3. Rahmani AM, Thanigaivelan NK, Gia TN, Granados J, Negash B, Liljeberg P, Tenhunen H (2015) Smart e-Health gateway: Bringing intelligence to Internet-of-Things based ubiquitous healthcare systems. In: 2015 12th annual IEEE consumer communications and networking conference, CCNC 2015, (June), 826–834. https://doi.org/10.1109/CCNC.2015.7158084
4. Say J, Tomasco MF, Heller A, Gal Y, Aria B, Heller E, Colman FC (2014) U.S. Patent No. 8,880,137. Washington, DC: U.S. Patent and Trademark Office

5. Schmidt AD, Schmidt HG, Batyuk L, Clausen JH, Camtepe SA, Albayrak S, Yildizli C (2009) Smartphone malware evolution revisited: android next target? In: 2009 4th international conference on malicious and unwanted software, MALWARE 2009. https://doi.org/10.1109/MALWARE.2009.5403026
6. Hutchins EM, Cloppert MJ, Amin RM (2011) Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Lead Issues Inf Warfare Secur Res 1(1):80
7. Wu DJ, Mao CH, Wei TE, Lee HM, Wu KP (2012) DroidMat: android malware detection through manifest and API calls tracing. In: Proceedings of the 2012 7th Asia joint conference on information security, AsiaJCIS 2012. https://doi.org/10.1109/AsiaJCIS.2012.18
8. ENISA threat landscape report 2017, Publications Office of the EU, ISSN: 2363-3050, https://doi.org/10.2824/967192
9. Faruki P, Bharmal A, Laxmi V, Ganmoor V, Gaur MS, Conti M, Rajarajan M (2015) Android security: a survey of issues, malware penetration, and defenses. IEEE Commun Surv Tutorials. https://doi.org/10.1109/COMST.2014.2386139
10. Zaslavsky A, Perera C, Georgakopoulos D (2013) Sensing as a service and big data. arXiv preprint arXiv:1301.0159
11. Burguera I, Zurutuza, U, Nadjm-Tehrani S (2011) Crowdroid: Behavior-based malware detection system for android. In: Proceedings of the 1st ACM workshop on security and privacy in smartphones and mobile devices—SPSM '11. https://doi.org/10.1145/2046614.2046619
12. Shabtai A, Kanonov U, Elovici Y, Glezer C, Weiss Y (2012) "Andromaly": a behavioral malware detection framework for android devices. J Intell Inf Syst. https://doi.org/10.1007/s10844-010-0148-x
13. Fei B, Eloff J, Olivier M, Venter H (2006) Analysis of web proxy logs. In: Olivier MS, Shenoi S (eds) Advances in digital forensics II. Springer, US, Boston, MA, pp 247–258
14. Amini S, Verhoeven R, Lukkien J, Chen S (2011) Toward a security model for a body sensor platform. In: 2011 IEEE international conference on consumer electronics (ICCE). IEEE, pp 143–144
15. Gajanayake R, Iannella R, Sahama TR (2012) Privacy oriented access control for electronic health records. In: Data usage management on the web workshop at the worldwide web conference. ACM
16. Garcia-Morchon O, Wehrle K (2010) Modular context-aware access control for medical sensor networks. In: Proceedings of the 15th ACM symposium on access control models and technologies. ACM, pp 129–138
17. Hoang VT, Krovetz T, Rogaway P (2015) Robust authenticated-encryption AEZ and the problem that it solves. In: Annual international conference on the theory and applications of cryptographic techniques. Springer, Berlin, pp 15–44
18. Maw H, Xiao H, Christianson B (2013) An adaptive access control model for medical data in wireless sensor networks. In: Proceedings of 2013 IEEE 15th international conference on e-Health networking, applications and services. IEEE
19. Ramli SN, Ahmad R, Abdollah MF, Dutkiewicz E (2013) A biometric-based security for data authentication in wireless body area network (wban). In: 2013 15th international conference on advanced communication technology (ICACT), IEEE, pp 998–1001

Part III
Applications, Data Mining and Data Analytics for E-Health

Chapter 8
DAS-Autism: A Rule-Based System to Diagnose Autism Within Multi-valued Logic

Saoussen Bel Hadj Kacem, Amel Borgi, and Sami Othman

Abstract In front of the continued growth of autistics number in the world, intelligent systems can be used by non-specialists such as educators or general physicians in autism screening. Moreover, it can assist psychiatrists in the diagnosis of autism to detect it as early as possible for early intervention. We propose in this chapter a tool for the diagnosis of autism: DAS-Autism. It is a knowledge-based system that handles qualitative knowledge in the multi-valued context. For this, we use our knowledge-based system shell RAMOLI, and its inference engine executes an approximate reasoning based on linguistic modifiers that we have introduced in a previous work. We have built a knowledge base that represents the domain expertise, in collaboration with a child psychiatry department of Razi hospital, the public psychiatric hospital in Tunisia. We have then conducted an experimental study in which we compared the system results to expert's diagnoses. The results of this study were very satisfactory and promising.

8.1 Introduction

Researchers are increasingly confronted with the need to support imperfect data in intelligent systems. In addition to the need to take account of this imperfection, one of the objectives of this work is to design systems that act as human behavior. Indeed, human mind uses imperfect knowledge that can be uncertain, vague, imprecise, etc.

S. Bel Hadj Kacem (✉)
Faculty of Economics and Management of Nabeul, Carthage University, Nabeul, Tunisia
e-mail: Saoussen.BelHadjKacem@ensi-uma.tn

COSMOS, National School of Computer Sciences, Manouba University, Manouba, Tunisia

A. Borgi
Institut Supérieur d'Informatique, LIPAH, Université de Tunis-El Manar, Tunis, Tunisia
e-mail: Amel.Borgi@insat.rnu.tn

S. Othman
Faculty of Medicine of Tunis, Tunis-El Manar University, Child Psychiatry Department, Razi Hospital, Manouba, Tunisia

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_8

Fuzzy logic was introduced by Zadeh [1] to handle these types of knowledge. It has attracted a lot of interest and has been used in several fields of application [2–5]. However, according to some authors [6–8], using fuzzy logic to represent abstract terms from natural language, such as *ugly*, *beautiful*, *intelligent*, is complicated and artificial to realize. Indeed, in fuzzy logic context, every term is modeled by a fuzzy set, which is based on a numerical domain. Nevertheless, abstract terms do not refer to numerical scales, which makes their modeling with fuzzy logic difficult and artificial. Symbolic multi-valued logic [6–9] is another logic that allows a symbolic representation of terms, and it is based on multi-set theory.

We have built in [10] a knowledge-based system shell in the multi-valued framework, called Raisonnement Approximatif basé sur les MOdificateurs LInguistiques (RAMOLI). It includes a data manager to introduce symbolic knowledge and an inference engine to reason with this knowledge. The inference engine uses an approximate reasoning based on linguistic modifiers that we proposed in [11, 12]. We propose in this work to use our approach of approximate reasoning through a practical application. Thus, our goal is to construct a knowledge-based system using RAMOLI and then evaluate its performance.

Knowledge-based systems are used in various fields to solve various problems: diagnosis (diseases, failures, etc.), decision on treatment, prognosis, etc. [13–16]. RAMOLI works in the context of multi-valued logic. All data are therefore represented symbolically. Its use is thus advantageous in applications which handle symbolic data. This is the case of medicine. Each medical specialty has its own way to define and establish diagnosis. It is done by collecting symptoms drawn from the patient's state. A symptom may have either a numeric value or a symbolic value. The doctor obtains symptoms with numerical values through measuring equipments (blood pressure monitor, glucose meter, blood test, etc.), whereas for symbolic information, he refers to the patient interrogation or to his own observation.

We chose in this work to build a knowledge-based system for medical diagnosis, more precisely for autism diagnosis. We chose psychiatry because its symptoms are mostly qualitative. Thus, it is easy to represent and manipulate them in our work environment.

Cases of autism are increasing worldwide. Zablotsky et al. [17] estimate that in 2016, 2.76% of American children are autistic, against 2% in 2012 [18] and 1.16% in 2007 [19]. The most serious problem is that the diagnosis is often made too late. This is because there are not enough specialists knowledgeable about various ways in which autism can appear [20]. This consequently causes a delay in the treatment of autism. Thus, it is necessary to detect autism as early as possible for early intervention. The aim of our work, as we initiated it in [21], is to help non-specialists such as educators or general physicians in autism screening. Also it will allow assisting psychiatrists in the diagnosis of autism. We call this system diagnosis aid system of autism (DAS-Autism) [22].

This chapter is organized as follows. In Sect. 8.2, we present autism, and we specify the limits of the tools already proposed in the literature for the diagnosis of this disease. Section 8.3 is devoted to knowledge representation in symbolic way, and we briefly describe the basic concepts of multi-valued logic, the context of our work.

We also present our approximate reasoning which is based on linguistic modifiers. Then in Sect. 8.4, we describe the design of the knowledge base and the development of the knowledge-based system DAS-Autism. An experimental study is presented in Sect. 8.5, for that we use real cases and compare the system results to experts diagnoses. Finally, Sect. 8.6 concludes the work.

8.2 Autism Diagnosis

We begin in this section by presenting autism. We also cite some diagnosis systems for autism that we found in the literature.

8.2.1 Description of the Domain

Autism usually begins in the early years of childhood (before age three) [23]. It is defined as a pervasive developmental disorder (PDD), which is characterized by severe development alterations in three areas [24]:

- Verbal and nonverbal communication;
- Social interaction;
- Behavior, interests and activities that are restricted and stereotyped.

Specialists make the diagnosis by observing the behavior of the patient and by questioning parents, referring to some standard protocols. The most widely used manual in psychiatry is diagnostic and statistical manual of mental disorders (DSM-IV-TR) [24], Published by the American Psychiatric Association (APA). It provides diagnostic criteria and classifications of mental disorders. We can also cite the ICD-10 (International Statistical Classification of Diseases and Related Health Problems) [25], published by the World Health Organization (WHO). It is a medical disease classification that includes a chapter devoted to mental and behavioral disorders.

Specialists also use evaluation questionnaires designed for the diagnosis of autism. They specify the intensity at which a child is autistic and allow monitoring of the disorders evolution. These questionnaires do not provide an entirely correct assessment in all cases, but help physicians to validate their opinions and to detect fragility signs of the child. The most used assessment instrument of autism diagnosis is childhood autism rating scale (CARS) [26]. It determines if a child is autistic and assesses the severity of the syndrome. It is a questionnaire of 14 symptoms of autism. For each symptom, a score is assigned on a scale expressing its severity. Despite its performance, CARS does not meet all the criteria of DSM-IV-TR and ICD-10. Autism Diagnostic Interview-Revised (ADI-R) [27] is another questionnaire. It is a tool of semi-structured interview with parents. ADI-R is based on the diagnostic criteria of DSM-IV-TR, ICD-10 and the latest knowledge in autism. It determines the diagnosis

with a threshold obtained by an algorithm. However, this instrument takes a long time (at least two hours).

In our work, we choose to refer to DSM-IV-TR. Indeed, this manual provides a fairly detailed description of autism and gives an algorithm for diagnosis aid. We also use CARS for the description of some symptoms. We will detail this idea later in this chapter.

8.2.2 Autism Diagnosis Tools

In computer science, researchers continue to make intelligent systems for medicine [28, 29], using various technologies: such as neural networks, genetic algorithms…, or combinations of these techniques. Some researchers were interested in some psychiatric diseases [30–33].

In the literature, some systems focus on autism diagnosis [20, 34–37]. These systems are based on data mining. Cohen et al. [20] use neural networks to differentiate children with autism and children with mental retardation. For this, 128 cases were used for learning and ten cases for the test. The average classification of the system is 92%. In the work of Arthi and Tamilarasi [34], the authors use fuzzy neural networks. Forty cases (patients) were collected for learning and for testing. The performance of this model is between 85 and 90%. Sunsirikul and Achalakul [35] use association-based classification to find behavior models for autistic and children with pervasive developmental disorder not otherwise specified. The clinical data in this study correspond to 140 patients and are operated by cross-validation. The average rate of correct classification is 85.27%. Kannappan et al. [36, 37] enforced the technique of fuzzy cognitive maps (FCM) on 40 cases. They had an accuracy percentage of 89.41%.

The disadvantage of neural network techniques is that they do not provide explanation of the diagnosis result. On the other hand, knowledge-based systems are able to give a trace of reasoning from inputs, triggered rules and chaining to attend the deduced result. Thus, the user can have an idea on the process covered by the system to achieve the provided result.

Another disadvantage of the systems described above is that they detect autistic children in a group which does not have a variety of associated pathologies. For example, the cases used in the system of Sunsirikul and Achalakul [35] are either autistic or suffering from pervasive developmental disorders. However, a system of autism diagnosis should be able to distinguish between autistic and normal children and children with other mental disorders. Indeed, some disorders have similar symptoms to those of autism.

8.3 Symbolic Knowledge-Based System

The context of our work is the multi-valued logic. We present in this section how to represent and manipulate imprecise knowledge to construct multi-valued knowledge-based systems in that context.

8.3.1 Knowledge Representation

Multi-valued logic introduces symbolic truth degrees which are intermediate between true and false [9]. According to this logic, every linguistic term is modeled by a multi-set. It generalizes classic set theory: The notion of belonging or not to a set is replaced by a partial belonging to a multi-set. The set of possible truth degrees is $\mathcal{L}_M = \{\tau_0, \ldots,\tau_i, \ldots,\tau_{M-1}\}$[1] with the total order relation: $\tau_i \leq \tau_j \Leftrightarrow i \leq j$, its smallest element is τ_0 (false), and the greatest is τ_{M-1} (true) [9, 38]. A possible list of truth degrees for $M = 7$ is $\mathcal{L}_7 = \{$*not-at-all, very-mildly, mildly, mildly-to-moderately, moderately, moderately-to-severely, severely*$\}$.

On the scale of truth degrees $\mathcal{L}_M$, operators can be defined to aggregate degrees as implications, T-norms and T-conorms. In multi-valued logic, the aggregation functions of Lukasiewicz are often preferred [9, 39].

These qualitative degrees can be considered as membership degrees of multi-sets. Indeed, "*X is* $\upsilon_\alpha A$" means that υ_α is the degree to which X satisfies the multi-set A.[2] In other words, the predicate A is satisfiable to a certain degree expressed through the scalar adverb υ_α associated with the truth degree τ_α of $\mathcal{L}_M$.

Multi-valued logic is based on the following interpretation:

$$\begin{aligned} X \text{ is } \upsilon_\alpha A &\Leftrightarrow X \text{ is } \upsilon_\alpha A \text{ is true} \\ &\Leftrightarrow \text{“}X \text{ is } A\text{”} \text{ is } \tau_\alpha\text{-true} \end{aligned}$$

For example, the statement "John is rather tall" means that John satisfies the predicate *tall* with the degree *rather*.

8.3.2 Approximate Reasoning Based on Linguistic Modifiers

In order to manage imperfect knowledge in intelligent systems, Zadeh has introduced the concept of *approximate reasoning* [40]. It is based on a generalization of *modus ponens* (MP) known as *generalized modus ponens* (GMP). This rule can be expressed in its standard form as follows:

[1] With M a positive integer not null, which represents the number of truth-degrees in the scale $\mathcal{L}_M$.

[2] Denoted mathematically by "$X \in A$": the object X belongs with a degree to the multi-set A.

$$\frac{\begin{array}{l}\text{If } X \text{ is } A \text{ then } Y \text{ is } B \\ X \text{ is } A'\end{array}}{Y \text{ is } B'} \tag{8.1}$$

where X and Y are linguistic variables and A, A', B and B' are fuzzy sets. GMP serves to infer not only with an observation exactly equal to the rule premise ("X is A"), but also with an observation which is different but approximately equal to it ("X is A'"). This allows handling imprecise knowledge in the inference process.

To determine the inference conclusion ("Y is B'"), a set of axioms is taken into account in order to have a logical and coherent result in concordance with human reasoning [41, 42]. In [11], we have proposed the generalization (8.2) of criteria appeared in [42]:

$$\left|\begin{array}{ll}\textbf{C I} & A' = A \Rightarrow B' = B \\ \textbf{C II-1} & A' \text{ is a reinforcement of } A \Rightarrow \text{ the more } A' \text{ is a reinforcement of } A \\ & \text{the more } B' \text{ is a reinforcement of } B \\ \textbf{C II-2} & A' \text{ is a reinforcement of } A \Rightarrow B' = B \\ \textbf{C III} & A' \text{ is a weakening of } A \Rightarrow \text{ the more } A' \text{ is a weakening of } A \\ & \text{the more } B' \text{ is a weakening of } B\end{array}\right| \tag{8.2}$$

Existing works in multi-valued framework of Akdag et al. [9] do not respect these axioms (see [11]). We introduced in a previous work [11, 43] an approximate reasoning that checks this axiomatics more precisely, criteria I, II-1 and III. These criteria allow having a gradual reasoning, which is adequate for our application of autism. Indeed, the severity of autism is proportional to the severity of the observed symptoms.

The proposed approximate reasoning is based on linguistic modifiers. A linguistic modifier is a function that expresses the modification that a predicate must undergo to become another predicate. In the multi-valued framework, modification of predicates is performed by dilation or erosion of the scales, and/or increasing or decreasing of the truth degrees. Akdag et al. [44] introduced linguistic modifiers in the multi-valued context and called them *generalized symbolic modifiers*. An example of these modifiers is the conserved reinforcing (CR) operator, which reinforces the degree by ρ and conserves the base:

$$CR_\rho = \begin{cases} \tau_{i'} = \tau_{min(i+\rho, M-1)} \\ \mathcal{L}_{M'} = \mathcal{L}_M \end{cases}$$

with ρ is the radius. The GMP of our approximate reasoning based on linguistic modifiers is the following:

$$\frac{\begin{array}{l}\text{If } X \text{ is } \upsilon_\alpha A \quad \text{then } Y \text{ is } \upsilon_\alpha B \\ X \text{ is } m(\upsilon_\alpha A)\end{array}}{Y \text{ is } m(\upsilon_\alpha B)} \tag{8.3}$$

where X and Y are linguistic variables, A and B are multi-sets, and υ_α and υ_β are linguistic degrees associated to the truth degrees υ_α and υ_β in $\mathcal{L}_M$. For the GMP (8.3), the observation is modeled by a modification of the rule premise $m(\upsilon_\alpha A)$, where m represents a linguistic [44].

In addition to check axiomatics, our approximate reasoning is very advantageous when knowledge is qualitative. Indeed, this type of knowledge is represented and managed easily by symbolic multi-valued logic. We recall that in fuzzy logic, knowledge, even the qualitative ones, is modeled by fuzzy sets. They are represented by a fuzzy membership functions on a numerical and continuous universe. So, reasoning with qualitative knowledge in fuzzy logic necessitates a complex matrix calculation. However, this is avoided with our approximate reasoning.

We extended this approximate reasoning in [45] to handle with heterogeneous knowledge. We mean by this heterogeneity that the multi-set in the observation is not necessarily the same as that of the rule premise, and/or the multi-set in the inferred conclusion is not necessarily the same as that of the rule conclusion. This offers more flexibility in the inference process.

Moreover, sometimes expert knowledge must be modeled by complex rules, i.e., rules whose premises are conjunction or disjunction of propositions. For this reason, we improved our approximate reasoning in [12] to deal with complex rules. We introduced for that new operators that aggregate linguistic modifiers: M-norm and M-conorm. M-norm, denoted by A_T, allows aggregating linguistic modifiers in a conjunction of propositions and is associated to a T-norm T. M-conorm, denoted by A_S, is for the disjunction case and is associated to a T-conorm S. We have proved that these aggregators verify logical connectives properties. For example, the aggregation of two modifiers $CR_{\rho 1}$ and $CR_{\rho 2}$ for the conjunction case is

$$A_T(\mathrm{CR}_{\rho 1}, \mathrm{CR}_{\rho 2}) = \mathrm{CR}_{\rho 3} \text{with } \tau_{\rho 3} = T(\tau_{\rho 1}, \tau_{\rho 2})$$

8.3.3 Knowledge-Based System Shell

A knowledge-based system shell is a generic tool that allows the construction of knowledge-based systems. It provides a software platform for building a knowledge base and provides a generic inference engine that allows the deduction of new knowledge.

In previous work [10], we have developed a knowledge-based system shell for symbolic multi-valued knowledge, we called it RAMOLI. This shell is a generic tool

that can be used in any field. Domain expertise is represented by multi-valued production rules and facts by multi-valued propositions. The inference engine implements exact reasoning as well as approximate reasoning. We have integrated our approximate reasoning based on linguistic modifiers that we had proposed in [11, 12, 43, 45]. The system is interactive and has GUIs that allow introducing knowledge base and triggering inference engine.

The construction of a knowledge base in RAMOLI is made by a set of stages. First, the user must introduce the basis that he will use, with their sizes and linguistic degrees terms. After, he specifies the manipulated multi-sets and associates to each of them a base. He must also add the linguistic variables that he will use. Once done, facts and rules can be constructed by doing combinations of linguistic variables, multi-sets and truth degrees. The inference engine can be executed after the filling of the knowledge base. It performs a forward chaining while considering imprecision and adds new facts to the knowledge base.

RAMOLI was developed in Java programming language. Thus, it provides platform portability, extensibility and easy integration with other Java code or applications.

8.4 Construction of DAS-Autism

Our goal is to build a knowledge-based system to aid diagnosis of autism: DAS-Autism. In what follows, we explain the steps that we followed for the development of this system: design of the knowledge base and system implementation.

8.4.1 Design of the Knowledge Base

Knowledge acquisition consists in acquiring knowledge from experts and in formalizing it. The formalization requires identifying involved concepts. These concepts are represented in our system by predicates, while expert knowledge will be represented by rules. We chose the formalism of rules because, in the autism domain, expert knowledge is easily translated into rules.

In this stage, our goal is to determine the set of predicates and rules that represent expert knowledge about autism. This is done by interviewing experts of the considered domain. However, in medicine, particularly in psychiatry, diagnosis strategies may vary from a doctor to another. Thus, the expertise of the interviewed doctors will influence the system result. System performance will not only depend on our approaches of inferences implementation, but more strongly on the quality of the knowledge acquisition phase, i.e., the involved doctors' expertise and its translation to a knowledge base.

Initially, to facilitate knowledge acquisition, our work is based on DSM-IV-TR [24]. Indeed, it provides an algorithm for decision making. An extract of the algorithm is given in Appendix A. From this algorithm, we have first extracted autism symptoms, and we have defined their linguistic variables and predicates in order to represent them. Then we built our rules base using these symptoms. But the symptoms shown in the DSM-IV are not nuanced, i.e., only their presence or absence in children is considered. However, our discussions with the psychiatrists have shown that they nuanced the symptoms. We recall that in our RAMOLI system predicates are multi-valued and can have several degrees. We therefore enriched extracted predicates by assigning a scale of ordered symbolic degrees. In a second step, we validated this knowledge base by domain experts, namely child psychiatrists from Razi hospital, the public psychiatric hospital in Tunisia.

We have extracted a total of 15 symptoms from the DSM-IV-TR algorithm. Identified symptoms are shown in Table 8.1. We note that autism may manifest many other symptoms. But these last symptoms may not characterize autism or are correlated with other symptoms. Their integration into the diagnosis is delicate, without warranty of betterment.

The used multi-sets for symptoms are *impaired* and *present*. Indeed, autism manifests some characters which are not present in a normal child. For these symptoms, we use the multi-set *present* to express the degree of presence of this character. Other characters are present in a normal child, and their presence in an autistic occurs in an altered way. For this reason, we use for these symptoms the multi-set *impaired*. As in the case of CARS, intensity of identified symptoms is qualified by degrees belonging to a scale of seven degrees:

Table 8.1 Symptoms list of autism

1	Nonverbal behaviors are impaired
2	Ability to develop peer relationships is impaired
3	Willingness to share is impaired
4	Social reciprocity is impaired
5	Emotional reciprocity is impaired
6	Development of spoken language is impaired
7	Conversation is impaired
8	Stereotyped or idiosyncratic language is present
9	Imitative play is impaired
10	Stereotyped patterns of interest are present
11	Patterns of interest abnormal in intensity are present
12	Patterns of interest abnormal in focus are present
13	Inflexible or ritual, non-functional adherence is present
14	Stereotyped and repetitive motor mannerisms are present
15	Persistent preoccupation with parts of objects is present

$\mathcal{L}_7 = \{\upsilon_0 = \text{not-at-all},$
$\upsilon_1 = \text{very-mildly},$
$\upsilon_2 = \text{mildly},$
$\upsilon_3 = \text{mildly-to-moderately},$
$\upsilon_4 = \text{moderately},$
$\upsilon_5 = \text{moderately-to-severely},$
$\upsilon_6 = \text{severely}\}.$

Recall that the system is intended not only to psychiatrists, but also to non-specialists. To have accurate and efficient facts, it is necessary to provide the most help to the user without taking into account their pre-knowledge about the disease.

In the CARS questionnaire, a manual is provided to the user. It explains each symptom and its degrees. It indicates the status of the child for whom a degree is chosen.

Similarly, we have associated with each degree of each symptom an explanation of the patient status. Some of these explanations are extracted from CARS, and others are provided from child psychiatrists of Razi hospital.[3] We give the user an explanation of degrees υ_0, υ_2, υ_4 and υ_6 of the base $\mathcal{L}_7$. The other degrees are considered intermediate degrees between these latter.

The final result is also in the form of a multi-set. The objective is the diagnosis of autism, so the chosen multi-set is *autistic*. It is represented by a base of four degrees:

$\mathcal{L}_4 = \{\upsilon_0 = \text{not-at-all},$
$\upsilon_1 = \text{mildly},$
$\upsilon_2 = \text{moderately},$
$\upsilon_3 = \text{severely}\}.$

Thus, the diagnosis result can have various degrees. Thus, the system not only indicates whether the child is autistic or not. As do child psychiatrists, it is able to indicate the severity of impairment of the child with the disease. We have also integrated other symbolic predicates, which are global symptoms deducted from symptoms entered by the user, such as *social interaction* or *communication*. These predicates will be used in chaining process.

Once the predicates are defined, the next step is to build the rule base which represents expert knowledge. We used for this the diagnosis algorithm of autism in the DSM-IV-TR [24]. For each symptom, we associated a rule whose premise is that symptom. Other rules were added and used for the deduction of intermediate global symptoms. Our rule base comprises a total of 23 rules. We give in appendix an extract from the rule base.

[3] Our set of symptoms, which is extracted from the DSM-IV, is not equivalent to that of CARS. Common symptoms between DSM-IV-TR and CARS are listed in Table 8.1 with the numbers 1, 3, 5, 6, 9, 13, 14 and 15.

Fig. 8.1 Principal menu of DAS-Autism

8.4.2 Development of DAS-Autism

In order to provide an appropriate system for autism diagnosis, we have developed specific GUI. Packages dedicated to the generation of knowledge-based systems in RAMOLI are included in DAS-Autism. Moreover, the necessary knowledge for the construction of the knowledge base and their characteristics, namely linguistic variables, predicates, scales degrees and rules are already introduced in the source code. The main window of the system is shown in Fig. 8.1.

"Open a diagnosis" and "List of diagnoses" buttons allow viewing the diagnoses already registered in the system. The "Option" button allows the setting of the KBS. It gives the possibility to change the T-norm and/or the T-conorm used for aggregation of modifiers.

When the user creates a new diagnosis, a first window appears, to enter information about the child. Then, a chain of windows occurs successively, each one for a symptom. Each interface provides the ability to choose the level of the corresponding symptoms. Figure 8.2 shows as an example the interface of the symptom *Willingness to share*. As we can see in the figure, the symptom name and its degrees have help buttons at the left. These buttons give descriptions and explanations in order to help the user in choosing the appropriate intensity of the symptom. For example, the description of *moderately altered* is shown in Fig. 8.3.

At the end of the questionnaire, a window displaying the result of the diagnosis appears (see Fig. 8.4). It specifies the severity of impairment of the child, i.e., the intensity of autism. Similarly, a button provides the ability to see the trace of reasoning. The trace specifies all the steps taken to arrive, from the introduced symptoms, to the deducted result of diagnosis.

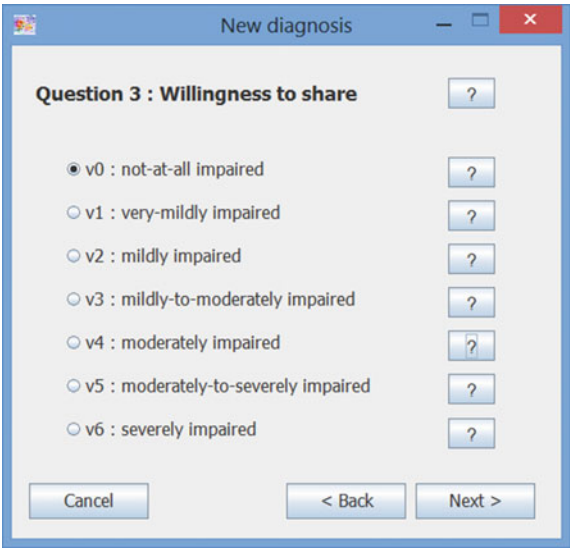


Fig. 8.2 Window of the symptom "Willingness to share"

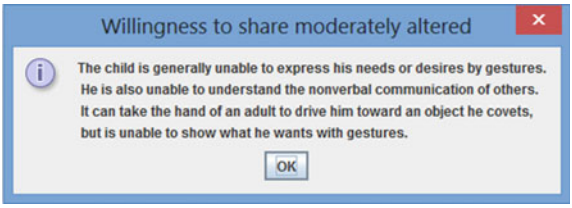


Fig. 8.3 Description dialog of "Willingness to share moderately altered"

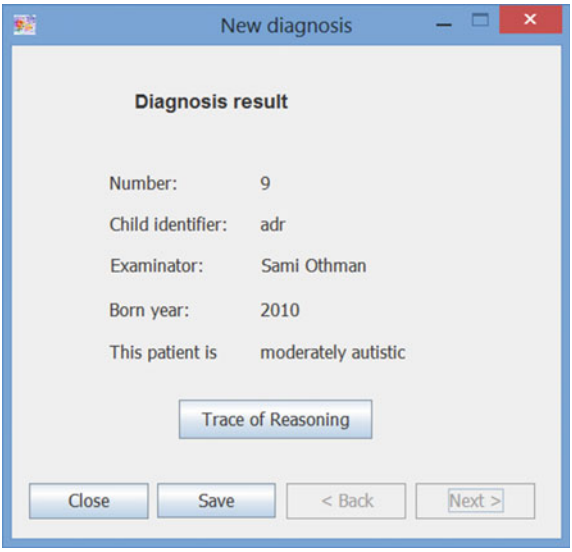


Fig. 8.4 Window of diagnosis result

8.5 Experimental Study

After constructing the DAS-Autism system, it is important to evaluate it. The objective is to determine whether it is effective enough to make valid diagnoses. To do so, the system result is compared with the diagnosis of expert to verify if they agree on their diagnoses and on the degree associated with their diagnosis. We also consider positive diagnosis for non-autistic cases (i.e., false-positive cases) and negative diagnoses for autism (i.e., false-negative cases) in order to detect potential errors of the system.

There is no common data set or benchmark for the diagnosis of autism. Every related work (see Sect. 8.2.2) has used its own testing set, which does not have the same attributes as the other works. In order to perform this experimental study, we used real cases with associated diagnoses of experts from child psychiatry department of Razi hospital. The given diagnoses are the synthesis of agreement after meeting and discussion of the child psychiatrists. The service has provided us a set of 40 cases [22]. Among these cases, 29 are autistic, with various degrees of severity. These cases allow us to see if the system gives true-positive diagnoses and/or false-negative diagnoses. Similarly, they allow comparing the severity degrees of system diagnoses with those given by the experts. The other 11 cases correspond to a set of children who are not autistic, but contain healthy children and other patients with other pathologies such as depression, mental retardation and infantile psychosis. This set is considered to check if the system provides true-negative diagnoses and/or false-positive diagnoses. Table 8.2 shows the partitioning of the test set according to the decisions of experts.

DAS-Autism has two parameters: T-norm and T-conorm. They are used in approximate reasoning to aggregate linguistic modifiers with M-norm and M-conorm. In this experimental study, we chose to use the T-norm and T-conorm of Zadeh *min* and *max*.

Table 8.3 shows the results of this study. We noticed from this experimental study that it does not provide false-positive and false-negative diagnoses. Therefore, the correct classification rate (CCR) is 100%. We then compared the severity degree of diagnoses. We found that for 29 autistics, the system gives the same severity degree as the expert for 23 cases. For the remaining six cases, the difference of severity degrees is of a unit for each case. The CCR becomes then 85% when considering

Table 8.2 Partitioning of the test set

Intensity	Cases number
Not-at-all autistic	11
Mildly autistic	7
Moderately autistic	15
Severely autistic	7
Total	40

Table 8.3 Result of the experimental study of DAS-Autism

Result	Cases number
True-positive	29
True-negative	11
False-positive	0
False-negative	0
Total	**40**
Correct severity degree	34
Incorrect severity degree	6 (gap = 1 degree)
Total	**40**

classification degrees, with a mean squared error (MSE) of 0.02. Therefore, the overall performance of the system is estimated at 93%.

These results are very satisfactory and promising. Immediate perspective would be to continue the experimental tests. It would be more interesting to compare our results with those obtained by other systems in the literature on common databases. For now, if we compare other systems to DAS-Autism (Sect. 2.2), we see that DAS-Autism provides better results. This comparison is of course to be qualified to the extent that the test base is different.

8.6 Conclusion

Imperfection becomes an inherent aspect of knowledge in knowledge-based systems. Its management allows getting as close as possible to the opinion of the expert. In this context, we chose to use symbolic multi-valued logic to handle such type of knowledge. We have proposed in this chapter a symbolic knowledge-based system for the diagnosis of autism, called DAS-Autism. We were based for the construction of this system on a knowledge-based system shell for symbolic multi-valued data, called RAMOLI [10]. More precisely, we used a package that allows introducing knowledge (rules and symptoms) and to perform an inference engine for deducing new facts (the diagnosis). We also implemented a specific GUI for this application to make easier the symptoms' entry. Then, a knowledge acquisition was necessary in order to model the expertise of the autism diagnosis. For that, we built a rule base in collaboration with psychiatrists in Razi hospital, and we were also based on DSM and CARS. We finally conducted an experimental study of DAS-Autism with real cases from Razi hospital. The obtained results are very satisfactory and enable a first validation of our work, both practically and theoretically. In this particular context, our approximate reasoning provides good results. The next step will be to deploy our diagnosis aid tool of autism among general practitioners.

Acknowledgements Authors would like to thank the psychiatrist Dr. Zied Choubani and the child psychiatrists Pr. Mohamed Bechir Hlaiem and Pr. Ahlem Belhadj for their help to elaborate this

work. They particularly opened for us the child psychiatry service of Razi hospital and provided us access to anonymous patient records.

Appendix 1. Extract from the Diagnosis Algorithm of DSM-IV-TR

A. A total of six (or more) items from (1), (2) and (3), with at least two from (1), and one each from (2) and (3):

   1. Qualitative impairment in social interaction, as manifested by at least two of the following:
      a. Marked impairment in the use of multiple nonverbal behaviors such as eye to-eye gaze, facial expression, body postures and gestures to regulate social interaction;
      b. Failure to develop peer relationships appropriate to developmental level;
      c. A lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing or pointing out objects of interest);
      d. Lack of social or emotional reciprocity.
   2. Qualitative impairments in communication as manifested by at least one of the following:
      a. Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gestures or mime);
      b. In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others.
      c ….

Appendix 2. Extract from the Knowledge Base of DAS-Autism

A. If social interaction is impaired and communication is impaired and restricted/repetitive/stereotyped behavior is present, then patient is autistic

   1. Social interaction:
      a. If nonverbal behaviors are impaired, then social interaction is impaired.
      b. If ability to develop peer relationships is impaired, then social interaction is impaired.
      c. If willingness to share is impaired, then social interaction is impaired.
      d. If reciprocity is impaired, then social interaction is impaired.

   i. If social reciprocity is impaired, then reciprocity is impaired.
   ii. If emotional reciprocity is impaired, then reciprocity is impaired.
2. Communication:
   a. If development of spoken language is impaired, then communication is impaired.
   b. If conversation is impaired, then communication is impaired.
   c. …

References

1. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353
2. Adnan MRHM, Sarkheyli A, Zain AM, Haron H (2015) Fuzzy logic for modeling machining process: a review. Artif Intell Rev 43(3):345–379
3. Xiang X, Yu C, Lapierre L, Zhang J, Zhang Q (2017) Survey on fuzzylogic-based guidance and control of marine surface vehicles and underwater vehicles. Int J Fuzzy Syst
4. Liu W, Liao H (2017) A bibliometric analysis of fuzzy decision research during 1970–2015. Int J Fuzzy Syst 19(1):1–14
5. Clarence W (2018) de Silva. In: Intelligent control. CRC Press
6. De Glas M (1989) Knowledge representation in a fuzzy setting. Report 89–48, LAFORIA, University of Paris VI
7. Pacholczyk D (1992) Contribution au traitement logico-symbolique de la connaissance. PhD thesis, University of Paris VI
8. Chung H-T, Schwartz DG (1995) A resolution-based system for symbolic approximate reasoning. Int J Approx Reasoning 13(3):201–246
9. Akdag H, De Glas M, Pacholczyk D (1992) A qualitative theory of uncertainty. Fundam Inform 17(4):333–362
10. Kacem SBH, Borgi A, Tagina M (2013) Ramoli: a generic knowledge based systems shell for symbolic data. In: World congress on computer and information technology (WCCIT), pp 1–6, Sousse, Tunisia
11. Kacem SBH, Borgi A, Ghédira K (2008) Generalized modus ponens based on linguistic modifiers in a symbolic multi-valued framework. In: Proceeding of the 38th IEEE international symposium on multiple-valued logic, pp 150–155, Dallas, USA
12. Kacem SBH, Borgi A, Tagina M (2015) Extended symbolic approximate reasoning based on linguistic modifiers. Knowl Inf Syst 42(3):633–661
13. Balakrishnan K, Honavar V (2011) Intelligent diagnosis systems. J Intell Syst 8(3–4):239–290
14. Sanchez Pi N, Carbo J, Molina JM (2012) A knowledge-based system approach for a context-aware system. Knowl Based Syst 27:1–17
15. Zhang Yi, Chen H, Jie Lu, Zhang G (2017) Detecting and predicting the topic change of knowledge-based systems: a topic-based bibliometric analysis from 1991 to 2016. Knowl Based Syst 133:255–268
16. Abu-Nasser BS, Abu Naser SS (2018) Rule-based system for watermelon diseases and treatment. Int J Acad Inf Syst Res (IJAISR) 2(7):1–7
17. Zablotsky B, Black LI, Blumberg SJ (2017) Estimated prevalence of children with diagnosed developmental disabilities in the United States, 2014–2016. NCHS Data Brief (291):1–8
18. Blumberg SJ, Bramlett MD, Kogan MD, Schieve LA, Jones JR, Lu MC (2013) Changes in prevalence of parent-reported autism spectrum disorder in school-aged us children: 2007 to 2011–2012. Nat Health Stat Rep 65(20):1–7
19. Kogan MD, Blumberg SJ, Schieve LA, Boyle CA, Perrin JM, Ghandour RM, Singh GK, Strickland BB, Trevathan E, van Dyck PC (2009) Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the us, 2007. Pediatrics 124(5):1395–1403

20. Cohen IL, Sudhalter V, Landon-Jimenez D, Keogh M (1993) A neural network approach to the classification of autism. J Autism Dev Disord 23:443–466
21. Kacem SBH, Borgi A, Othman S (2016) A diagnosis aid system of autism in a multi-valued framework. In: Uncertainty modelling in knowledge engineering and decision making (FLINS 2016), pp 405–410, Roubaix, France
22. Kacem SBH (20130 Un raisonnement approximatif basé sur les modificateurs linguistiques et son intégration dans les systèmes à base de connaissances symboliques multi-valents. PhD thesis, National School of Computer Sciences, University of Manouba
23. Dumas JE (2007) Psychopathologie de l'enfant et de l'adolescent. Ouvertures psychologiques. De Boeck, 3rd edn
24. American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders DSM-IV-TR Fourth Edition (Text Revision). American Psychiatric Publishing, Washington, DC, 4th edn
25. World Health Organization (1993) International statistical classification of diseases and health related problems ICD-10. World Health Organization, Geneva, 10th edn
26. Schopler E, Reichler R, DeVellis R, Daly K (1980) Toward objective classification of childhood autism: Childhood autism rating scale (cars). J Autism Dev Disord 10:91–103
27. Lord C, Rutter M, Couteur A (1994) Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24:659–685
28. Pandey B, Mishra RB (2009) Knowledge and intelligent computing system in medicine. Comput Biol Med 39(3):215–230
29. Mahfouf M, Abbod MF, Linkens DA (2001) A survey of fuzzy logic monitoring and control utilisation in medicine. Artif Intell Med 21(1–3):27–42
30. Pluggea LA, Verheya FRJ, Jollesa J (1990) A desktop expert system for the differential diagnosis of dementia: an evaluation study. Int J Technol Assess Health Care 6:147–156
31. Bichindaritz I (1994) A case-based assistant for clinical psychiatry expertise. In: Proceedings of the annual symposium on computer application in medical care, pp 673–677
32. Chattopadhyay S, Pratihar DK, De Sarkar SC (2008) Developing fuzzy classifiers to predict the chance of occurrence of adult psychoses. Knowl Based Syst 21(6):479–497
33. Chattopadhyay S, Pratihar D (2010) Towards developing intelligent autonomous systems in psychiatry: its present state and future possibilities. In: Pratihar D, Jain L (eds) Intelligent autonomous systems, volume 275 of studies in computational intelligence, pp 143–166. Springer, Berlin
34. Arthi K, Tamilarasi A (2008) Prediction of autistic disorder using neuro fuzzy system by applying ann technique. Int J Dev Neurosci 26(7):699–704
35. Sunsirikul S, Achalakul T (2010) Associative classification mining in the behavior study of autism spectrum disorder. In: The 2nd international conference on computer and automation engineering (ICCAE'10), pp 279–283, Singapore
36. Kannappan A, Tamilarasi A, Papageorgiou EI (2011) Analyzing the performance of fuzzy cognitive maps with non-linear hebbian learning algorithm in predicting autistic disorder. Expert Syst Appl 38(3):1282–1292
37. Papageorgiou EI, Kannappan A (2012) Fuzzy cognitive map ensemble learning paradigm to solve classification problems: application to autism identification. Appl Soft Comput 12(12):3798–3809
38. Ginsberg ML (1988) Multivalued logics: a uniform approach to reasoning in artificial intelligence. Comput Intell 4(3):265–316
39. Bartusek T, Navara M (2001) Conjunctions of many-valued criteria. In: Proceedings of the international conference uncertainty modelling, Bratislava, Slovakia, pp 67–77
40. Zadeh LA (1975) The concept of a linguistic variable and its application to approximate reasoning—i—ii—iii. Inf Sci 8:199–249, 8:301–357, 9:43–80
41. Baldwin JF, Pilsworth BW (1980) Axiomatic approach to implication for approximate reasoning with fuzzy logic. Fuzzy Sets Syst 3(2):193–219

42. Fukami S, Mizumoto M, Tanaka K (1980) Some considerations of fuzzy conditional inference. Fuzzy Sets Syst 4(3):243–273
43. Borgi A, Kacem SBH, Ghédira K (2008) Approximate reasoning in a symbolic multi-valued framework. In: Lee RY, Kim HK (eds) Computer and information science, volume 131 of studies in computational intelligence, pp 203–217. Springer
44. Akdag H, Truck I, Borgi A, Mellouli N (2001) Linguistic modifiers in a symbolic framework. Int J Uncertainty Fuzziness Knowl Based Syst 9(Supplement):49–61
45. Kacemv SBH, Borgi A, Tagina M (2009) On some properties of generalized symbolic modifiers and their role in symbolic approximate reasoning. In: ICIC'09, volume 5755 of lecture notes in computer science, pp 190–208. Springer, Berlin

Chapter 9
Smart E-Health Home Supervision Systems

Eric Campo, Damien Brulin, Daniel Estève, and Marie Chan

Abstract Supervisory systems have become interesting solutions for monitoring the health of home-based people. Frail older people are at high risk of becoming dependent if they are not cared for quickly. Embedded devices at home or on the person are two possible options which, when coupled, allow a more precise knowledge of behaviors and a faster triggering of the alert. The idea still applied is to model the person's activities or health parameters in a continuous and automated way to detect deviations from the usual behavior. This clinico-technical approach has been developed and tested in different places and by several research teams with a real medical interest in having a longitudinal knowledge of a patient's behavior. However, prevention and follow-up of home prescriptions must go beyond the purely technological and medical aspects and take economical and organizational dimensions into account. We consider that three types of complementary actions must be implemented: a fundamental action on the scientific and technological level to know how to build, from the data collected, "a profile of user activity;" detecting, in real time, all kinds of deviations from the normal pattern; training and ground demonstrations actions.

D. Estève—deceased.

E. Campo (✉) · D. Brulin · D. Estève · M. Chan
LAAS-CNRS UPR8001, Université de Toulouse, CNRS, UT2J, Toulouse, France
e-mail: campo@laas.fr

D. Brulin
e-mail: dbrulin@laas.fr

D. Estève
e-mail: esteve@laas.fr

M. Chan
e-mail: chan@laas.fr

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_9

9.1 Introduction

In the previous chapter, we saw that the knowledge acquisition was necessary to model the expertise of a disease diagnosis. This modeling is based on rules and algorithms building for decision making. This can assist psychiatrists diagnose autism in order to detect it as early as possible for early intervention. In smart home applications for health monitoring, signal acquisition follows the same learning approach to get knowledge about people's behavior.

The issue of home care patients' monitoring (convalescents, elderly, frail or disabled people…) has been identified for more than 20 years as a major societal challenge. Advances in medicine and living comfort mean that the elderly population is growing in number and proportion to the overall population [1]. The initial idea was to make technological progress benefits to the continuous monitoring of the elderly, firstly for demographic reasons (ageing populations) and because there is a strong demand for these people to maintain their independence as long as possible [2]. At the same time, recently, technologies have evolved favorably: connected sensors, data mining, analysis and decision support software are available to monitor people automatically and continuously. This idea, largely initiated with the goal of keeping elderly people at home, has gradually opened up to other patient profiles: "frail", convalescent and disabled. This evolution is becoming possible thanks to the advent of telemedicine, the computerization of hospitals and the increasing systematization of personalized medical records [3].

Many R&D projects have been devoted to this issue worldwide, and so far, few of them have met with great success with users. It is, therefore, necessary to draw up an overview of existing technologies and to identify the obstacles that slow down their evaluation in order to propose, in the long term, new development and valuation options.

This chapter is organized as follows:

- a review of the technology currently available,
- a better identification and analysis of needs,
- identification of obstacles encountered in the field of health monitoring,
- the orientations of a renewed offer.

9.2 Historical Review of the Technology Currently proposed

The issue of elderly observation "in general" goes back to the 1990s, with growing awareness of accelerated ageing of populations in the most developed countries. The US Census Office predicts that in Europe, by 2050, 30% of the population will be over 65 [4]. Since the 2000s, major studies have been initiated by the European Union Member States participating in the EURHOMAP [5] project to reflect the

"Home Care" situation in all countries of the community. In our societies, it is clearly necessary to meet the three major requirements of patients wishing to:

- stay independent, at home, as much as possible,
- access to strong health and safety support,
- enjoy comfort and conviviality, especially for people living alone.

In France, in 1996, the CNRS mandated our team to set up a multidisciplinary working group to assess the impact of computerization on housing. This assessment concluded that there was an urgency need to address the home care of the elderly [6]. Many authors, from all disciplines, became aware of this need, at this time and made proposals to provide at least partial solutions [7–9]. These solutions have been based primarily on care assessment approaches and on the influence of physical activity interventions on health and psychological functions. We want here to consider, first of all, the technical and technological solutions that have been proposed to "monitor and secure the elderly" for almost 15 years [10]. Since then, technologies, particularly Information and Communication Technologies (ICT), have rapidly evolved to allow, today, an industrial exploitation of the initial concepts. Our thinking is based on these evolutions, in the form of "connected and distributed sensors" which ensure the close monitoring of people and surrounding risks, the collection and storage of data, the follow-up ensuring the processing and, possibly, the emission of alarms to the actors of surveillance [11]. We have realized that the technological option that we are going to research does not address the whole problem. Other key sociocultural issues need be taken into consideration [12]:

- questions of acceptability for use,
- privacy issues,
- social, family and human ethical issues.

With all these dimensions, helping the elderly and patients living at home, in general, is a very difficult problem for which we will only consider the following points:

- *Ensuring the safety of the elderly*: They may be chronically ill or suffer after accidental falls, physical or mental failures, discomfort or illness. It is then up to the "automatic" surveillance to quickly detect any risk situation and to inform relatives via a local or remote "intelligence" (24 hours a day) that will manage and propose the most appropriate intervention.
- *Longitudinal monitoring*: Home care should not be limited to the emergency detection of a personal risk; the need for assistance and health and safety monitoring of daily life must be met and thus the data from automatic surveillance must be combined with the data, interventions and decisions of other organizations that already exist or will exist (home assistance devices, reinforcement of remote surveillance, connections to medical records and medical specifiers…).
- *Conviviality*: This is a strong requirement to allow regular links with family and friends so that contact remains strong with social life.

Although these new technologies (ICT) can be used as a communication medium, this last point is not strictly technical and is outside our field of analysis. We will deal with the problem from the point of view of the instrumentation that can be implemented to secure people at home and ensure their daily monitoring. In this instrumentation approach, we identify three main options which are, in fact, steps in the design of home monitoring systems for older adults: home instrumentation alone; instrumentation of the person and the environment and networking of all those involved in monitoring.

9.2.1 Home Instrumentation

The first idea was to design a completely transparent home monitoring system to the patient, with the primary objective of not bothering him/her in any way and to fully respect his/her privacy. This general option converges towards an only instrumentation of the accomodation: observing and characterizing particular occupations of the patient [13] in order to ensure that they are carried out as usual, showing the normality of the situation. Technologically, this is a first-generation option based on the use of IR presence sensors [14], most recently applied to the use of cameras [15]. The sensors are distributed in the home environment at strategic points, connected to a wireless beacon, in a conventional continuous monitoring architecture. It is in this field that "smart home" demonstrations are multiplying, trying to meet all home automation needs [16].

The methodological developments related to the very particular surveillance of people are the most interesting: The first experiments quickly showed that patients, for the most part, reproduce their behavior "day after day". The idea then is to set up a model of these usual behaviors and to use them as a model to anticipate "normal" behavior. Then, it becomes possible, by comparing current data and data anticipated by the model, to continuously detect deviations from the usual behavior and, thus, to transmit alarms when the deviations are considered too high.

This approach has notably been investigated by the LAAS-CNRS, which has tested it both at home [17] and in institutions [18, 19]. It requires some definitions:

- "Usual behavior", for us, consists of the most reproducible elements of behavior, characteristic of a lifestyle, in a time slot, from one day to another, from one week to another. The finer the breakdown of the behavior, the more precise the anticipation model can be and, therefore, provide a more reliable alarm trigger.
- "Indicators" are the direct or aggregated data provided by the sensors; these can be the presence rate, schedule of identifiable activities, usual travel, motion speed and others.

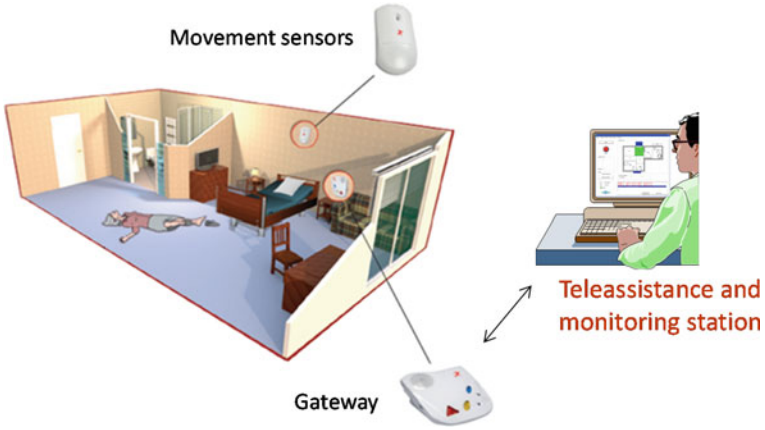


Fig. 9.1 Instrumented home for 24-h fall monitoring

This approach is now found in the literature [20–22], which has been validated at various test sites at home or in institutions and is now offered commercially by several small and medium-sized enterprises.[1]

This approach is well accepted because it reassures families and does not impose on patients. It can be usefully enhanced by the implementation of more sophisticated sensors (camera) or more specialized (safety sensors in bed, kitchen, bathroom…) and by other communication technologies: connected sensors and interface connecting the technical system, the patients and the surveillance actors [23]. This type of patient monitoring at home joins the more general problem of securing goods and people [24].

But it is an approach that has limits related to patient's behavior change that induces imprecision in behavior patterns: The consequences are regular errors of diagnosis [25]. In general, we evaluate the performance of a detection system by what are called false positives and false negatives:

- False negatives—the system did not detect the danger,
- False positives—the system has set off an unfounded alarm.

This difficulty is particularly troublesome in emergency situations: in an approach where only the home is instrumented (Fig. 9.1), the detection of a fall or faintness is built on a criterion of immobility, which is supposed to set the threshold detection beyond a time calculated by learning techniques [26]. This delay, due to the imprecision of the normality model, is necessarily important and often incompatible with the requirement to intervene, for example, within a maximum of 45 min following a cardiac or cerebral event [27].

[1] Téléassistance Senioradom, https://www.senioradom.com/
Telegrafik, Service Otono-me, https://www.telegrafik.fr/
Senioralerte, téléassistance active, http://www.senioralerte.com/
Domosafety, http://www.domo-safety.com/.

Fig. 9.2 Robot assistant in a life environment

Following the same logic, we can imagine the replacement of the presence sensor by cameras: the image can remain blurred to respect the patients' privacy, but, in case of an alert, the image can be used to visualize the situation and validate the risks [28]. This option is now credible because the use of cameras has become widely popular, especially for the protection of a house against intrusion. Habits also have evolved: Faced with the need for security, few people still seriously mention the right to total privacy.

Home instrumentation is therefore a simple, ready-to-use and inexpensive approach to securing people. Its effectiveness remains limited, with risks that can be controlled by a small involvement of the patient, for example, by providing a manually operable warning button. There is the possibility of having a robot assistant [29] for the patient at home, which could be, in case of an alert, programmed, among other functions, to visually check the state of the situation (Fig. 9.2).

9.2.2 Mixed Equipment: Patient at home

In the version that consists only of home instrumentation, the risk to the patient is related to the non-detection of faintness or a fall. Cameras do not see all the space of an indoor environment, and the immobility criterion can define a significant decision-making time, which is dangerous in an emergency. To reduce this risk, the perfectly identified priority need is embedded fall detection [30]. Thanks to miniaturization, current technology can meet this need with several low-cost integrated sensors: accelerometers, barometers, gyrometers, on the market today, are

different detection software proposals that combine three complementary criteria, as follows [31]:

- the drift of the patient's activity and behavior,
- sudden change in acceleration at the time of a fall,
- break of the activity before/after a presumed fall.

This is a very powerful solution that allows a quick decision and reduces false positives or false negatives to almost zero. The practical trend is to propose a bracelet [32] on which a warning button can be integrated. This solution can also be included as an additional option in the 24-hour continuous monitoring structure already described in Sect. 9.2. Obviously, the position of the wrist sensor is not the least risky solution, depending on the intensity of manual activity that can interfere with detection. Some suggest placing the sensor on a belt [33] and others recommend placing it in a shoe [34]. In the surveillance of dementia with Alzheimer's disease, it is proposed to place the sensor in the back to keep patients from tearing at the patches [35].

9.2.3 Tomorrow: "Patient" Involvement

It now can be seen that hardware and software tools are in place to provide effective systems for detecting any life incident. An interesting part of the research is moving towards the exploitation of continuously collected data, beyond alert situations that are treated as an emergency by specialized operators 24 hours a day. This exploitation can take two main directions:

- longitudinal behaviors follow up [36] to detect changes that would indicate a long-term risk, using technology for prevention,
- complement the action of surveillance by orienting it further towards the monitoring of specific physiological parameters [37].

These new steps cannot be conceived without the direct integration of technical monitoring into the practices of the medical world [3]. This assumes that the medical world appropriates the tool and its use for therapeutic purposes: prevention of health risks, diagnoses, follow-up prescriptions, convalescences and home care. The University Hospital of Toulouse, for example, proposes a pathway [38, 39] for the use of surveillance technologies in the monitoring and support for "frail people". The implementation consists of welcoming patients during regular visits for prescriptions and equipping them with wearable instrumentation to ensure that they follow these prescriptions. Useful feedback validates the recommandation and, if necessary, modify it with full knowledge of analyzed facts. In the example used, home monitoring is carried out according to Fried's criteria [40].

Technically, this extension of embedded measurements is boosted by sports tracking applications [41] based on immediate feedback to the user via his/her cell phone. This technical option completes the system architecture (Fig. 9.3) and invites the multiplication of embedded measurements to develop a real embedded computer

Fig. 9.3 Overview of integrated clinical and technical surveillance

network [42] communicating in real time with all the actors involved and the patient himself.

The accumulation of patient data opens two new ways:

- the data processing to make it useful for medical treatment. The aim is to develop synthetic, efficient and user- friendly interfacesfor all actors in the healthcare network, including patients,
- the extraction of these databases to advance knowledge and medical practice (integration of monitoring in the processing of medical records, statistics…) [43].

9.3 Further Analysis of Health Technological Needs

With the advent of "digital" technology, we should see major changes in the health care system. The aspects we have just mentioned are only the premises of this evolution, guided by the needs of patients, medical practice and the market. We will come back to the market which takes these new technologies, particularly home monitoring technologies, in chapter's conclusion. But, here we want to open two new challenges that should greatly benefit from the contributions of digital technology: the prevention of health risks and the follow-up at home of medical prescriptions.

9.3.1 Prevention

The typical example of current events is the monitoring of "Frailty," which is already the topic of important works. Frailty *is a common geriatric syndrome that embodies an elevated risk of catastrophic declines in health and function among older adults* (Wikipedia). Using the frailty phenotype framework proposed by Fried et al. [40], prevalence estimates of 7–16% have been reported in non-institutionalized,

community-dwelling older adults [44]. The occurrence of frailty increases progressively with age and the concerns of those of lower socioeconomic status. Frail older adults are at high risk of major adverse health effects, including disability, falls, institutionalization, hospitalization and mortality. Fried proposed five dimensions for assessing frailty:

- involuntary weight loss,
- exhaustion,
- muscle weakness,
- slowness of the walk and
- low levels of activity.

On these medical bases (indicators and prospects of care), we have designed a frailty monitoring at home largely based on the implementation of an instrumented sole [39]. This concept is supported by the fact that the wearing of an insole can be "forgotten" by the patient while it provides interesting information on gait and balance characteristics and serves as a particularly effective fall detector (Fig. 9.4).

For the first time, this project combines the interests of health network actors (University Hospital of Toulouse), industry (Actia) and a technological laboratory to design a complete set of continuous patient monitoring systems (Fig. 9.5).

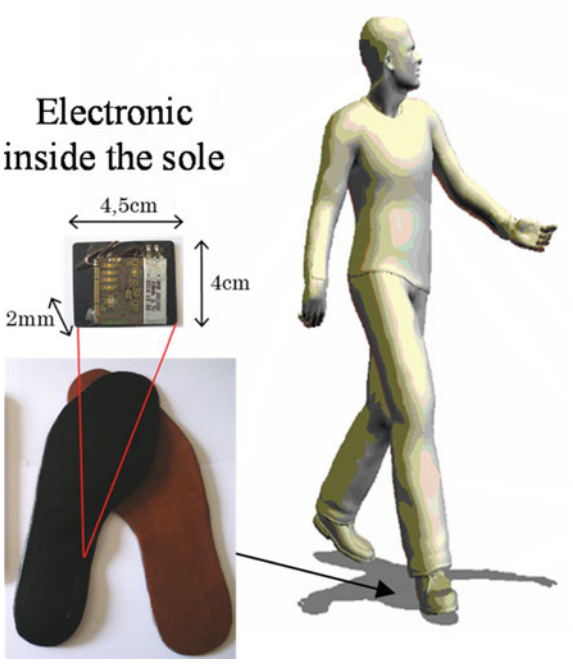


Fig. 9.4 Characteristics of the smart sole developed at LAAS-CNRS

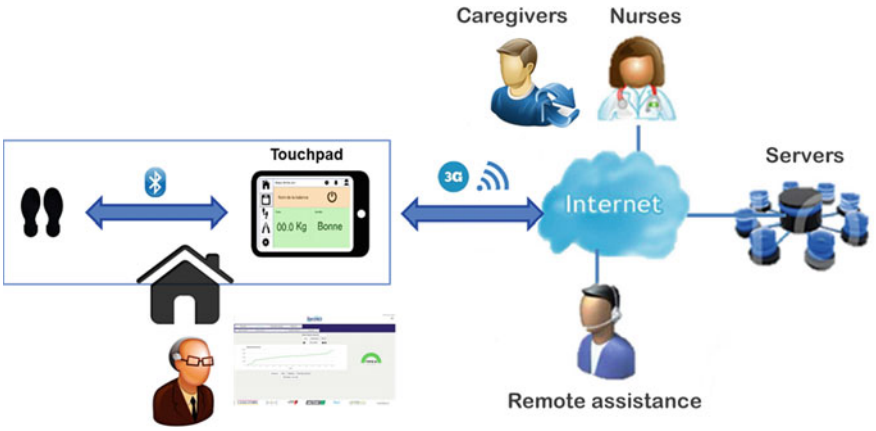


Fig. 9.5 Diagram of the FRAGIL'IT globale care structure

9.3.2 Prescription Follow-up

Telemonitoring methods and tools for "follow-up the implementation of medical prescriptions at home" define a general problem that we propose to address here, by following up, at home, the wearing of orthoses. This is, in itself, an important area that implies the use of a physical medium that can easily accommodate sensors and "intelligence" and that must lead quickly to medical and industrial valuations. The interest of this orientation also is to allow both the monitoring of the prescription and the patient's activity level. The doctor must be able to work with certainty on the drug intake or, in this case, the wearing of an orthosis. Moreover, patients' continuous monitoring is a source of new information that will make it possible, for example, to extract longitudinal follow-up, medium- and long-term alerts and prevention [45]. The objective is to provide additional information that cannot be observed in a medical practice or center. The new close relationship that needs to be established between the home-based patient and the attending physician is also likely to contribute solve the thorny issues of care organization related to the development of telemedicine, rural medicine, etc.

Figure 9.6 shows a schematic diagram of the design and development of an orthotic instrumentation using embedded and communicating electronic micromodules and data processing software, possibly merged with data from the "Patient Record," to define alarms. It also proposes a longitudinal risk prevention analysis. These two objectives open two fields of research as follows:

- the acquisition of personalized patient data concerning the use of orthopedic prescription, which implies the development of new embedded sensors and their connection to a collection network that will refer to a secure, accessible database, under condition, to the patient and all the actors of the health network,

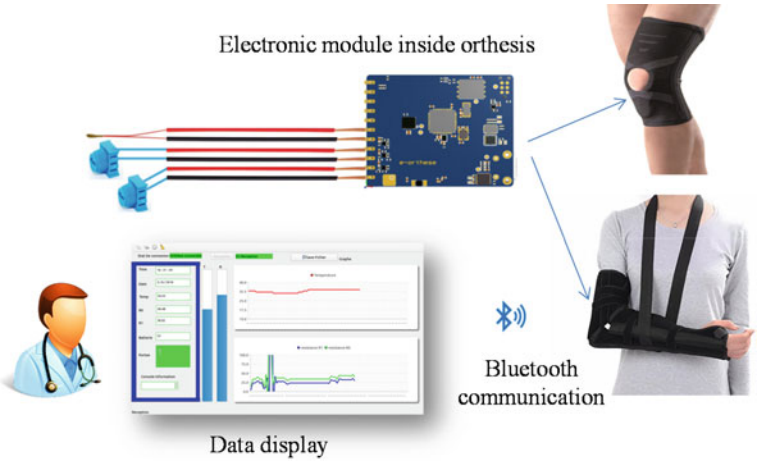


Fig. 9.6 Architecture for monitoring the wearing of orthoses at home

- processing of the collected data which includes filtering steps, learning models, merging prescription tracking data with other complementary data in the file or the patient's life context. This presupposes new developments in diagnosis and decision support algorithms.

9.4 Obstacles to Real-Life Implementation

The needs of patients, convalescents and the elderly people at home are well established: stay independent, at home, as much as possible; access to rapid and solid health and safety assistance; benefit from comfort and conviviality, especially for people living alone.

The proposed technologies makes it possible to meet these needs thanks to the hardware and software developments we have just described. There is no doubt that home monitoring is part of a general trend of exploiting technological advances to equip connected sensor systems and exploit the data collected in order to anticipate dangers and optimize treatments.

The obstacles to effective implementation are of two types: the non-existence of an economic model for home care and the organization of health services and social life.

9.4.1 Economic Component

In France, for example, the home monitoring sector is part of the dynamic market for medical devices: It is an annual market of about 15 billion euros (43% of the costs of these medical devices are reimbursed by health insurance, the rest by complementary health and households); it is the second-largest expense item after prescription drugs. An increase of this budget of 10% over ten years opens a potential investment of 1500 euros per installation on the basis of one million annual installations, covering 10 years of French society's needs. This objective is not unreasonable in that it can be shared between social security, complementary health and the family. Also, it can provide relief from current charges currently paid by families for supervision, medical assistance and, in some cases, placement in specialized facilities (nursing homes). Compared to the overall budget, France devotes a significant share of GDP to health expenditure (12% of GDP), which corresponds to 4000–5000 euros of annual expenditure per inhabitant, which is expected to increase by less than 2–3%. These simple evaluations allow us to see that the costs are reasonable but significant, which puts the implementation projects in a position of certainty with regard to effectiveness, patient and family acceptance and economic efficiency. What are the costs of organizational change? And what are the gains associated with this technological breakthrough?

In the French HOMECARE project, a market study realized by Atos[2] company would tend to show that home care is globally profitable for our societies. It is enough to compare the price of a day in hospital (400 euros) with that of a day in an institution (100 euros), to note that the technological monitoring investment of 1500 euros per installation is derisory: It can be largely amortized over one year in place. From all these economic considerations, it must be concluded that, faced with the immense problem of monitoring ageing, there are technological options that are or will soon be compatible with the economic development of our societies. The other slowing factors are of a socioeconomic and organizational nature.

9.4.2 Organizational Component

One of the specificities of a home care project is to set up a direct and/or an indirect communication architecture, putting the patient in contact with the family and medical environment for a progressive optimization of the "rendered benefit". The implementation leads to new practices for careers (patients' health data acquisition by the doctor without "traditional consultation," triggering an alert for the patient, the doctor or the nurse who comes every day to the home for the patient's follow-up, etc.). These new practices will therefore encourage new health care organizations to optimize the work of care teams that should help solve the difficult problems facing our health system: home support workers, medical deserts and telemonitoring.

[2] Atos Company, https://atos.net/.

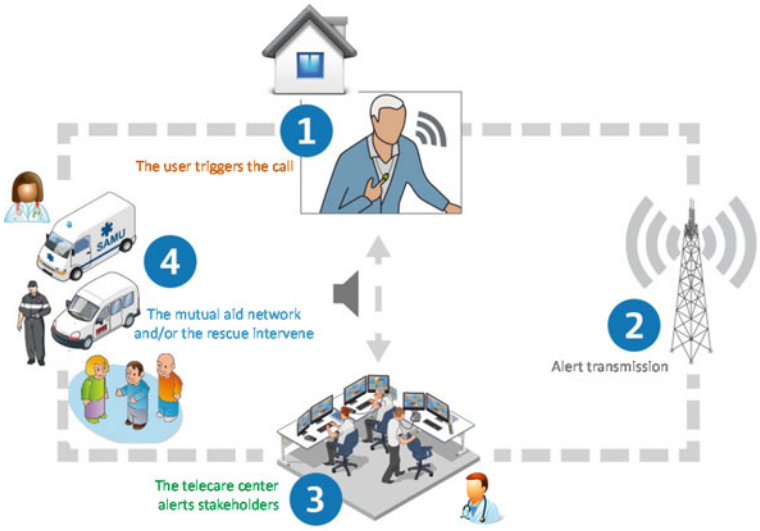


Fig. 9.7 Complete monitoring structure with central monitoring and intervention station

The project should also be articulated with all other existing forms of monitoring and intervention:

- There is already a societal care, more or less developed depending on the country which allows third parties to intervene at home for housework, bringing meals or simply making visits to maintain user-friendliness.
- There are also caregiver interventions for regular home care.
- Finally, in emergency cases, there are interventions by specialized organizations.

Figure 9.7 shows the basic structure of automated monitoring designed to provide 24-hour service: centralized stations (PCs) that bridge the gap between the patient's emergency risk and the interventions teams. They also provide daily follow-up by answering phone calls from worried patients. Early experiments show that there is a need for organizational work to articulate the hospital entries in case of emergency [46].

9.5 What Technological Offer for Tomorrow?

The urgent need to take into account the ageing of populations has prompted research into technological solutions for the surveillance, safety and comfort of elderly people living alone. These technological proposals have evolved rapidly since the 1990s, with the emergence of microtechnologies that promised very powerful monitoring instrumentation at low cost: the integration by analog device of the accelerometer dates back to 1991!

Historically, the technological response to needs has been the "warning button" [47]. Patients are instrumented with a necklace or wristband that must be used in case of emergency. This is an important advance because it has made it possible to structure a 24-hour monitoring service and to better organize the intervention of assistance staff and the response to emergencies. Practically, this solution is perfectly effective if the person being monitored is vigilant and scrupulously applies the instructions for use. However, studies on this topic [48] have shown that a large proportion of patients do not use the device permanently and that people suffering from dementia refuse it. Since 2001, review articles [49] have been published that pave the way for the home distributed instrumentation. The basis of the proposals remains the phone as a means of exchange between monitoring teams and patients. This development has led to the design and demonstration of the second-generation type of assistive equipment that can be used in three different situations:

- The approach incorporates the equipment into the community phone alarm system, with its ready-to-use infrastructure of communications, databases, control centers and response networks. The equipment is individually adapted to the needs of patients in their homes, based on the results of a "needs assessment."
- Unobtrusive sensors at home can be used to collect lifestyle pattern data. Lifestyle data is analyzed regularly and activity patterns recorded. A telephone message is automatically constructed and delivered to a caregiver.
- "Smart homes" incorporate numerous sensors, which are installed throughout the accommodation. The sensors are connected via a smart network to cut-off devices, which ensure the safety of the patient.

A third generation of "monitoring systems" is emerging in which sensors and actuators are integrated on humans to further minimize the risk of false alarms and to open new fields of application for frail or convalescent people [50]. In addition to the need for new highly specialized sensors and actuators [51], this evolution opens new needs for the development of a sensor–actuator network technology that can be embedded into humans [52] (Fig. 9.8).

Although all these contributions make an important contribution to the safety, security, independence and quality of life of elderly people living at home, there are some limitations.

In the obstacles encountered in the field, we have highlighted the problematic economic model, that is critical today. But the diffusion of these new technologies also asks many challenges as follows:

- Not everyone will benefit from or accept new technological aids and devices, and each individual's situation must be carefully assessed (customization of the technological offer).
- Many people may welcome the technology, although a few may consider it as invasion of privacy (respect for privacy).
- People with visual, hearing or speech disabilities may not be able to use some of the technology.

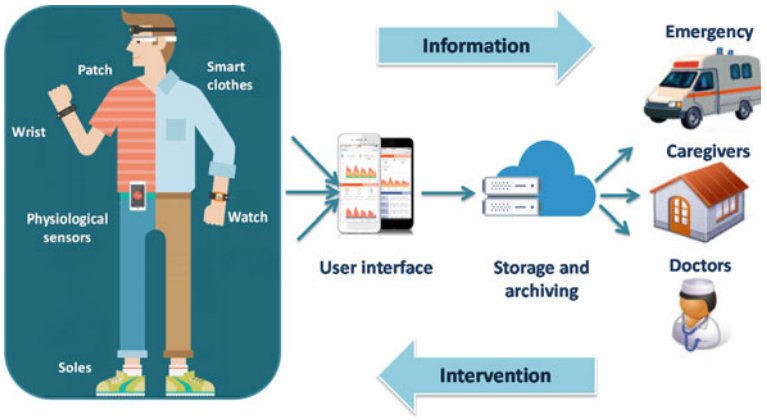


Fig. 9.8 BodyLan architecture

- We need to reassure the whole chain of people interested in the performance and reliability of the offer.
- The equipment is relatively too expensive.

To move forward in this context, we consider that three types of complementary actions need to be implemented:

1. **Fundamental action at the scientific and technological level**: The long-term emerging generic concept is how to build, from the collected data, "a profile of a user's activity and be able to detect, in real time, all kinds of deviations from the normal pattern."
   This implies important developments in sensors and embedded modules and especially significant advances in data processing and model building capable of taking into account the variability of behaviors.
2. **Training**: Progress must be integrated into existing surveillance networks, which may differ from one country to another.
   To do this, training campaigns are necessary to:

   - train staff, including home-based workers,
   - integrate the medical need,
   - involve families,
   - standardize technology to reduce costs.

   Such training courses should be both general (to explain the meaning of things) and specialized.
3. Field actions and demonstrations
   The interest of communities and insurance companies is to promote these technical solutions. Procedures and funding exist that should be more oriented towards experimentation with monitoring technologies:

- home help, which can be followed by regular visits of care assistants, service staff, charities. These aids are partly financed by the state or local authorities and constitute an important work in society with a need for better recognition.
- Specialized institutions are developing to welcome people most affected by ageing and who need daily care.

In the context of these existing procedures, the technological offer must be industrial and have an exemplary character in order to be evaluated. From the patients' point of view, it is necessary to adapt the offer to different situations, for a "personalized" monitoring, at least in cases where the person:

- is robust (self-control),
- is fragile (prevention),
- is at risk (24-hour surveillance),
- shows signs of dementia (24-hour supervision),

It is also necessary to find the right path between family involvement and medical practice. The connection with care structures—family physicians and hospital medicine—probably poses the most visible difficulties. Optimization work needs to be done on the clinical medicine-surveillance system interface, to assist physicians in their diagnosis and follow-up and to avoid "loading" physicians with useless data.

The strategy to be developed must be based on a generic architecture, which will be adapted to the type of patient and the patient him/herself, as well as to medical and family requirements.

9.6 Conclusion

This chapter very briefly summarizes 20 years of reflection work based on two questions: What technologies for health? How far will the impact of "digitization" affect the organization of the healthcare network?

These questions are important for two reasons:

- They reflect an urgent social and societal need, linked in particular to populations ageing and the strategy of keeping patients at home as much as possible.
- They fit perfectly into the development logic of digital technologies: connected embedded objects, database and digital monitoring.

The first observation we make in this analysis is that the context has changed considerably over the years between the 1990s and today:

- in terms of relationships between people to monitor and the methods and means proposed. Respect for the privacy of people was a major priority in the 1990s which has gradually been turned into a security and health priority,

- in terms of technologies developed specifically for each problem, towards more generic technologies carried by wireless mobile telephony and by the development of connected, miniaturized, distributed sensors, databases and monitoring software.

This evolution is visible in the technological progress we have presented:

- First, a direct instrumentation of housing that led to the concept of the "smart home" with, for the monitoring of the elderly living alone at home, the implementation of IR presence sensors. This step is based on the idea of basing risk detection on the modeling of habits that serve as a reference for usual behaviors. This idea has been favorably exploited on the basis of presence and displacement indicators, and is now a possible area for development.
- Home instrumentation alone has obvious limitations for people safety: The accuracy and relevance of alarms are too low. The ambition to be able to detect falls and discomfort has led to a first instrumentation of people by miniaturized "patches" that have at least three embedded functions: identification of the person, his/her location and detection of a fall.
- The stage that we are experiencing today is a new system step: In what specific application framework can a generic technology for monitoring people be developed? Several axes of development are open, which we have differentiated in this work according to the user characteristics:
  - For robust people, there is a dynamic of development around amateur or professional sport to monitor performance. The cell phone is an important element of this development, which allows the results of the monitoring to be read at any time.
  - For frail people, there is an emerging health and medical need with the main objective of monitoring home care. It is a complex system under development that involves all stakeholders: health care actors, field workers and the users.
  - For people suffering from a disability (physical or mental) or a chronic illness, the issue is autonomy and home care. In this chapter, we questioned the economic model and showed that the interest of such a development axis is widely shared between patients, their families, the medical world and society as a whole.

The results of this study undoubtedly show that the technologies are there, available to ensure the continuous monitoring of the elderly, but there is still a need to create omprehensive systems that meet the needs and to validate these systems so that the industrial commitment to operate the sector is made. There is still a step to be taken with the users who will contribute through experimentation to refine the proposals before they are fully satisfactory.

A strategic outcome of the continuous monitoring implementation is the availability of data over long observation periods. For the authors of the article, this data availability and even more knowledge accumulated by the system should allow for more refined decision making following the detection of a problem. For that,

the development of self-reconfigurable predictive models may possibly be a way to explore.

Health care network actors are faced with a probable upheaval in their practices, with increased means to prevent, correctly diagnose and monitor the evolution of patients' health.

The next chapter illustrates this predictive approach by observing data in order to use the generated patterns or the model in the area of cancer diagnosis.

Acknowledgements The authors thank the various colleagues and academic and industrial partners who participated in the projects mentioned in this chapter. Some of these projects were supported by the French National Research Agency (CNRS) and the Council of the Occitanie Region.

References

1. Department of Economic and Social Affairs, United Nations (2017) World Population Ageing—Highlights, New York
2. Sixsmith J, Sixsmith A, Malmgren Fänge A, Naumann D, Kucsera C, Tomsone S, Haak M, Dahlin-Ivanoffi S, Woolrych R (2014) Healthy ageing and home: the perspectives of very old people in five European countries. Soc Sci Med 106:1–9
3. Chan M, Campo E, Brulin D, Estève D (2017) Biomedical monitoring technologies and future healthcare systems. J Sci Technol 3(1):59–75
4. He W, Goodkind D, Kowal P (2016) An aging world: 2015. International population reports P95/16-1, March 2016
5. Genet N, Boerma WG, Kringos DS, Bouman A, Francke AL, Fagerström C, Melchiorre MG, Greco C, Devillé W (2011) Home care in Europe: a systematic literature review. BMC Health Serv Res 11(207)
6. Estève D (1997) Conclusions de la commission d'experts "habitat Intelligent". Rapport CNRS-SPI, Développement des Soins à Domicile
7. Merlis M (2000) Caring for the frail elderly: an international review. Health Aff 19(3):141–149
8. Stewart K, Challis D, Carpenter I, Dickinson E (1999) Assessment approaches for older people receiving social care: content and coverage. Int J Geriatr Psychiatry 14(2):147–156
9. McAuley E, Katula J (1998) Physical activity interventions in the elderly: influence on physical health and psychological function. Ann Rev Gerontol Geriatr 18(1):111–154
10. Alelaiwi A, Hassan MM, Bhuiyan MZA (2017) A secure and dependable connected smart home system for elderly. In:Dependable, autonomic and secure computing, 15th international conference on pervasive intelligence & computing, 3rd international conference on big data intelligence and computing and cyber science and technology congress, Orlando, FL, USA, 6–10 November 2017
11. Uddin M, Khaksar W, Torresen J (2018) Ambient sensors for elderly care and independent living: a survey. Sensors 18(7)
12. Zwijsen SA, Niemeijer AR, Hertogh CM (2011) Ethics of using assistive technology in the care for community-dwelling elderly people: an overview of the literature. Aging Mental Health 15(4):419–427
13. Korhonen I, Parkka J, Van Gils M (2003) Health monitoring in the home of the future. IEEE Eng Med Biol Mag 22(3):66–73
14. Suzuki R, Otake S, Izutsu T, Yoshida M, Iwaya T (2006) Monitoring daily living activities of elderly people in a nursing home using an infrared motion-detection system. Telemed J E Health 12(2):146–155

15. Fleck S, Straßer W (2008) Smart camera based monitoring system and its application to assisted living. Proc IEEE 96(10):1698–1714
16. Chan M, Estève D, Escriba C, Campo E (2008) A review of smart homes-present state and future challenges. Comput Methods Programs Biomed 91:55–81
17. Bonhomme S, Campo E, Estève D, Guennec J (2007) An extended PROSAFE platform for elderly monitoring at home. In: 29th annual international conference of the IEEE engineering in medicine and biology society (EMBEC 2007), Sciences and Technologies for Health, Lyon, France, 22–26 August 2007
18. Campo E, Bonhomme S, Estève D, Guennec J (2008) Remote monitoring of elderly in rural zones using multisensors system and satellite communication. In: International conference on smart homes and health telematics—ICOST2008, Gerontechnology: enhancing the quality of life for rural elders, Iowa State University, USA, 28 June–2 July 2008
19. Bourennane W, Charlon Y, Bettahar F, Chan M, Estève D, Campo E (2013) Ambient intelligence for monitoring Alzheimer patients. Int J E-Health Med Commun 4(1):23–35
20. Medjahed H, Istrate D, Boudy J, Dorizzi B (2009) Human activities of daily living recognition using fuzzy logic for elderly home monitoring. In: 2009 IEEE international conference on fuzzy systems, Jeju Island, South-Korea, 20–24 August 2009
21. Fleury A, Noury N, Vacher M (2011) Improving supervised classification of activities of daily living using prior knowledge. Int J E-Health Med Commun (IJEHMC) 2(1):17–34
22. Aztiria A, Izaguirre A, Augusto JC (2010) Learning patterns in ambient intelligence environments: a survey. Artif Intell Rev 34(1):35–51
23. Milenković A, Otto C, Jovanov E (2006) Wireless sensor networks for personal health monitoring: issues and an implementation. Comput Commun 29(13–14):2521–2533
24. Sarathy R (2006) Security and the global supply chain. Trans J 45(4):28–51
25. Taraldsen K, Askim T, Sletvold O, Einarsen EK, Grüner Bjåstad K, Indredavik B, Helbostad JL (2011) Evaluation of a body-worn sensor system to measure physical activity in older people with impaired function. Phys Ther 91(2):277–285
26. Noury N, Herve T, Rialle V, Virone G, Mercier E, Morey G, Moro A, Porcheron T (2000) Monitoring behavior in home using a smart fall sensor and position sensors. In: 1st annual international IEEE-EMBS special topic conference on microtechnologies in medicine and biologym Proceedings, Lyon, France, 12–14 October 2000
27. SMUR 15 (2013) Référentiel et guide d'évaluation en France. June 2013
28. De Miguel K, Brunete A, Hernando M, Gambao E (2017) Home camera-based fall detection system for the elderly. Sensors 17(12):2864
29. Korchut A, Szklener S, Abdelnour C, Tantinya N, Hernández-Farigola J, Ribes JC, Rejdak K (2017) Challenges for service robots—requirements of elderly adults with cognitive impairments. Front Neurol 8:228
30. Mubashir M, Shao L, Seed L (2013) A survey on fall detection: principles and approaches. Neurocomputing 100:144–152
31. Khan SS, Hoey J (2017) Review of fall detection techniques: a data availability perspective. Med Eng Phys 39:12–22
32. Nho Y, Lim JG, Kim D, Kwon D (2016) User-adaptive fall detection for patients using wristband. In: 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS), Daejeon, Korea, 9–14 Oct 2016
33. Bourke AK, van de Ven P, Gamble M, O'Connor R, Murphy K, Bogan E, McQuade E, Finucane P, Laighin GÓ, Nelson J (2010) Evaluation of waist-mounted tri-axial accelerometer based fall-detection algorithms during scripted and continuous unscripted activities. J Biomech 43(15):3051–3057
34. Sim SY, Jeon HS, Chung GS, Kim SK, Kwon SJ, Lee WK, Park KS (2011) Fall detection algorithm for the elderly using acceleration sensors on the shoes. In: 2011 annual international conference of the ieee engineering in medicine and biology society, Boston, MA, USA, 30 Aug–3 Sept 2011
35. Hajjine B, Escriba C, Medale D, Fourniols JY (2016) Design, integration and characterization of a tracking patch: application to elderly monitoring. J E-Health Telecommun Syst Netw 5(3):57–74

36. McAuley E, Morris KS, Motl RW, Hu L, Konopack JF, Elavsky S (2007) Long-term follow-up of physical activity behavior in older adults. Health Psychol 26(3):375
37. Pantelopoulos A, Bourbakis NG (2010) A survey on wearable sensor-based systems for health monitoring and prognosis. IEEE Trans Syst Man Cybern Part C (Appl Rev) 40(1):1–12
38. Piau A, Charlon Y, Campo E, Vellas B, Nourhashemi F (2015) A smart insole to promote healthy aging for frail elderly individuals: specifications, design, and preliminary results. JMIR Rehabil Assist Technol 2(1)
39. Charlon Y, Campo E, Brulin D (2018) Design and evaluation of a smart insole: application for continuous monitoring of frail people at home. Expert Syst Appl 95:57–71
40. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, McBurnie MA (2001) Frailty in older adults: evidence for a phenotype. J Gerontol Ser A: Biol Sci Med Sci 56(3):146–156
41. Tamm T, Pärlin K, Tiimus T, Leemets K, Terasmaa T, Must I (2014) Smart insole sensors for sports and rehabilitation. Proceedings of SPIE—the international society for optical engineering 9060
42. Patel M, Wang J (2010) Applications, challenges, and prospective in emerging body area networking technologies. IEEE Wirel Commun 17(1):80–88
43. Kvedar J, Coye MJ, Everett W (2014) Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff 33(2):194–199
44. Siriwardhana DD, Hardoon S, Rait G, Weerasinghe MC, Walters KR (2018) Prevalence of frailty and prefrailty among community-dwelling older adults in low-income and middle-income countries: a systematic review and meta-analysis. BMJ open 8(3)
45. Barlow J, Singh D, Bayer S, Curry R (2007) A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions. J Telemed Telecare 13(4):172–179
46. Scott WR, Ruef M, Mendel PJ, Caronna CA (2000) Institutional change and healthcare organizations: from professional dominance to managed care. University of Chicago Press, Chicago
47. Dibner AS, Lowy L, Morris JN (1982) Usage and acceptance of an emergency alarm system by the frail elderly. Gerontol 22(6):538–539
48. White EB, Montgomery P (2014) Electronic tracking for people with dementia: an exploratory study of the ethical issues experienced by carers in making decisions about usage. Dementia 13(2):216–232
49. Miskelly FG (2001) Assistive technology in elderly care. Age Ageing 30(6):455–458
50. Majumder S, Aghayi E, Noferesti M, Memarzadeh-Tehran H, Mondal T, Pang Z, Deen MJ (2017) Smart homes for elderly healthcare—recent advances and research challenges. Sensors 17(11):2496
51. Chen M, Gonzalez S, Vasilakos A, Cao H, Leung VC (2011) Body area networks: a survey. Mob Netw Appl 16(2):171–193
52. Talbi Y (2018) Conception d'un patch transdermique intelligent pour le monitoring et l'aide à la prise de médicament, Doctoral dissertation, INSA de Toulouse

Chapter 10
Literature Review: Overview of Cancer Treatment and Prediction Approaches Based on Machine Learning

Ahmed Maalel and Mahbouba Hattab

Abstract The purpose of this chapter is to provide a literature survey through an overview of the research fields relevant to cancer treatment and prediction approaches based on machine learning. The past few years have witnessed an exponential growth in databases and repositories due to the increase in scientific knowledge and the massive data production. Biomedical domain represents one of the rich data domains. An extensive amount of biomedical data is currently available with wealth of information, ranging from details of clinical symptoms to various types of biochemical data and outputs of imaging devices El Houby (J Appl Biomed, 2018 [1]). One of the important biomedical research domains, epidemiological cancer research, who is of high priority across the world. Cancer has been characterized as a heterogeneous disease consisting of many different subtypes. Every sixth death in the world is due to cancer, making it the second-leading cause of death (second only to cardiovascular diseases) Schutte (Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study, 2017 [2]). The early diagnosis and prognosis of a cancer type have become a necessity in cancer research, as it can facilitate the subsequent clinical management of patients. The vast amount of hidden data in huge databases related to cancer with all his variability has created tremendous interests in the field of data mining. While data mining is a discipline resulting from the combination of classical statistics and computer science algorithms, such as machine learning, aim to the extraction of new and useful knowledge from a large amount of data, it has become a useful instrument in bioinformatics. It can depict variation of cancer incidence and mortality by region, ethnicity, gender and socioeconomic factors that contribute to the assessment of population health needs, while it can contribute to the study of cancer burden. Furthermore, in-depth analysis of the patient's profile using data mining methods may uncover hidden, previously unknown relations between patient profile, cancer

A. Maalel (✉)
National School of Computer Sciences, RIADI Laboratory, University of Manouba, 2010 Manouba, Tunisia
e-mail: ahmed.maalel@ensi.rnu.tn

A. Maalel · M. Hattab
Higher Institute of Applied Science and Technology, University of Sousse, 4003 Sousse, Tunisia

H. Idoudi and T. Val (eds.), *Smart Systems for E-Health*, Advanced Information and Knowledge Processing, https://doi.org/10.1007/978-3-030-14939-0_10

treatment and surveillance. This chapter will present recent approaches and works related to this context. It presents also a synthesis of current and future trends for smart systems for E-health.

10.1 Introduction

During the past few years, the increase in scientific knowledge and the massive data production have caused an exponential growth in databases and repositories. The knowledge discovery in databases, the ability to extract useful hidden knowledge and the development of methods and techniques for making use of data are becoming increasingly important in today's competitive world. Biomedical domain represents one of the rich data domains. An extensive amount of biomedical data is currently available with wealth of information (see Fig. 10.1), ranging from details of clinical symptoms to various types of biochemical data and outputs of imaging devices [1]. The amounts of data generated by healthcare transactions are too complex and huge to be processed and analyzed by traditional methods. As a result, the science of data management and analysis is also advancing to enable organizations to convert this vast resource into information and knowledge that helps them achieve their objectives [3]. One of the very known domains dealing with this challenge is health informatics, who represent the field of information science concerned with the analysis, use and

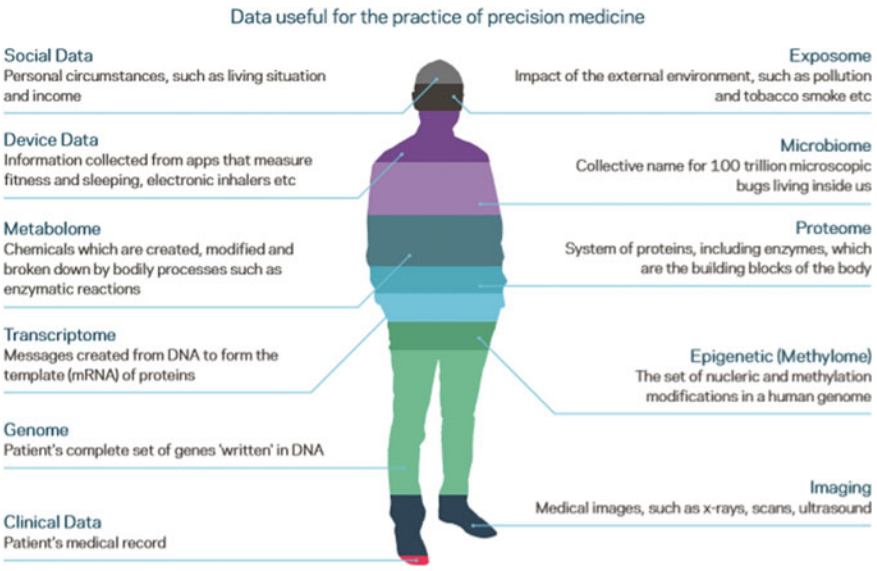


Fig. 10.1 Data useful for the practice of precision medicine (Introduction to machine learning in health care: http://web.orionhealth.com/)

dissemination of medical data and information through the application of computers to various aspects of health care and medicine[1].

From their side, biomedical researchers face also the same problem of finding important knowledge from this huge amount of data [1]. So, health informatics is a rapidly growing field that is concerned with applying computer science and information technology to medical and health data [4, 5]. Computing technologies and health informatics have been widely applied in medicine especially in critical situation such as cancer which is one of the important biomedical research domains and of high priority across the world.

Cancers are defined by the National Cancer Institute (NCI)[2] as a collection of diseases in which abnormal cells can divide and spread to nearby tissue. As this definition suggests, cancers can arise in many parts of the body and in some cases spread to other parts of the body through the blood and lymph systems. Cancer is the second-leading cause of death globally, next only to heart disease, in both developed as well as developing countries, and was responsible for 8.8 million deaths in 2015. Globally, nearly one in six deaths is due to cancer. In 2016, 8.9 million people are estimated to have died from various forms of cancer.[3] The Institute for Health Metrics and Evaluation (IHME) put relatively small error margins around this global figure: the lower and upper estimates extend from 8.75 to 9.1 million (Full data on cancer deaths, including upper and lower estimates can be downloaded at the IHME's Global Burden of Disease (GBD)). The overall drop in cancer death rates is largely due to decreasing death rates for lung, breast, prostate and colorectal cancers.[4]

Also shown in the charts below, Fig. 10.2 is total number of people suffering from cancer at any given time, differentiated by cancer type. This is measured across both sexes and all ages. The prevalence of breast cancer is the highest globally; an estimated 8 million had breast cancer in 2016; 6.3 million had colon and rectum; 5.7 million with prostate; and over 2.8 million with tracheal, bronchus and lung cancer [6].

The early diagnosis and prognosis of a cancer make it more likely to respond to effective treatment and can result in a greater probability of surviving, less morbidity and less expensive treatment. Significant improvements can be made in the lives of cancer patients by detecting cancer early and avoiding delays in care. The amount of data coming from instrumental and clinical analysis of this disease is quite large and builds a huge and wealthy dataset. Machine learning in medicine has recently made headlines and by applying machine learning methods to cancer databases and, with in-depth analysis of the patient's profile using statistical or data mining methods may uncover hidden, previously unknown relations between patient profile, cancer treatment and surveillance and consequently help solving diagnostic and prognostic problems related to cancer [7].

---

[1] National Library of Medicine, 2017.

[2] https://www.cancer.gov/.

[3] World Health Organization, February 2018.

[4] Facts and Figures 2018: Rate of Deaths from Cancer Continues Decline.

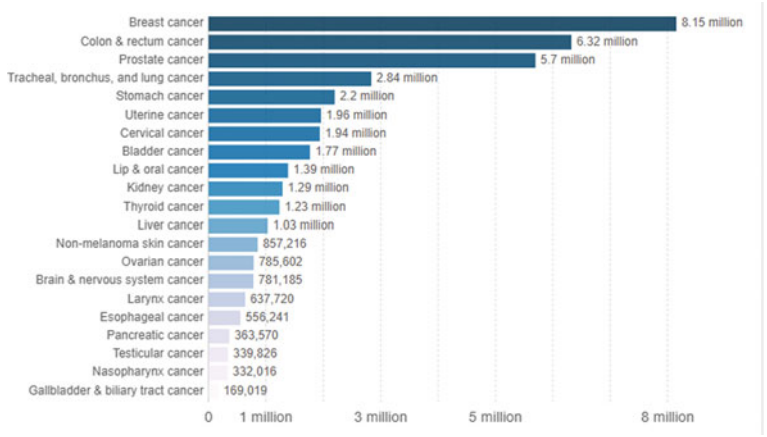


Fig. 10.2 Number of people with cancer by type, World, 2016 (IHME, Global Burden of Disease (GBD))

In the rest of this chapter, we will present the machine learning concepts as well as an overview of studies of the literature.

10.2 Machine Learning

Machine learning is to train the system over a large observed data to predict outcomes or categorize observations in future data, by other means, to generate extraction patterns or build a model and use the generated patterns or model to make predictions in the future for unknown cases. The dataset used to learn the model is known as the training dataset. The records making up the training set are referred to as training samples and are randomly selected from the sample population. The resulting model is typically applied to new samples to categorize or predict values of the outcome for previously unseen observations, and its performance evaluated by comparing predicted values to actual values for a set of test samples, this step is known as supervised learning [8]. Machine learning techniques are used to analyze important clinical parameters and their combinations for prognosis in cancer research domain, e.g., prediction of disease progression, extraction of medical knowledge for outcome research, therapy planning and support and for the overall patient management [9]. While data mining is a discipline resulting from the combination of classical statistics and computer science algorithms, such as machine learning, aim to the extraction of new and useful knowledge from a large amount of data, it has become a useful instrument in health informatics. It can depict variation of cancer incidence and mortality by region, ethnicity, gender and socioeconomic factors that contribute to the assessment of population health needs, while it can contribute to the study of cancer burden. Furthermore, in-depth analysis of the patient's profile using data

mining methods may uncover hidden, previously unknown relations between patient profile, cancer treatment and surveillance. Therefore, mining of medical data is one of the most interesting, though hard, machine learning tasks. Thus, both the data mining and healthcare industry have emerged some of reliable early detection systems and other various healthcare-related systems from the clinical and diagnosis data related to cancer in order to predict the different stages of the latter as well as to assisting in diagnosis for the doctors for making their clinical decision [10]. As we see, it seems the healthcare environment is becoming more and more reliant on computing technologies. The use of ML methods and data mining techniques can provide useful aids to assist physicians in many cases of cancer provide rapid identification of abnormalities and enable diagnosis in real time, considering the importance of time in detection and diagnosis of cancer. There are two categories of data mining models as it is shown below in Fig. 10.3: predictive model and descriptive model [12]. Predictive data mining often applies supervised learning functions to predict unknown or future values of other variables of interest [12]. Descriptive data mining often applies the unsupervised learning functions in finding patterns describing the data that can be interpreted humans [12]. An overview of the applications of predictive and descriptive data mining on the management of cancers is presented in the next section of this chapter. The included studies are classified according to data mining model applied on datasets.

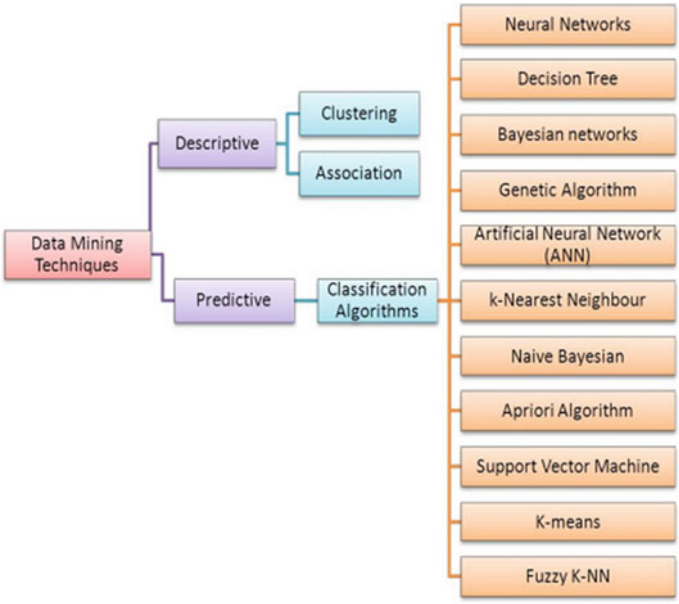


Fig. 10.3 Different techniques in healthcare domain [11]

10.3 An Overview of Studies of the Literature

In recent decades, due to the dramatically and exponential growth of cancer incidence and related deaths worldwide, significant progress has been made in the development of data mining. A variety of data mining techniques and approaches are applied to analyze and summarize data acquired from cancer databases, thereby extracting relevant knowledge. In this section, we review several data mining applications in the area of cancer diagnosis and prediction. At first, five study (approaches) are presented in the next section whose common point is the use of unsupervised learning to extract new knowledge.

Chauhan et al. [13]

In this study, the researchers used the data analytical tools and data mining techniques to cluster medical spatial datasets, which includes discovery and extraction of hidden, interesting and useful patterns from this large database by grouping the objects into clusters. This study focuses on discrete and continuous spatial cancer databases on which clustering techniques are applied and the efficient clusters were formed [13]. In case of continuous data, the clusters of arbitrary shapes are formed. Moreover, this application explored data mining techniques such as classical clustering and hierarchical clustering on the spatial cancer datasets to generate the relevant clusters. This study uses two different clinical databases, to determine relevant pattern detection for cancer diagnosis. The first clinical dataset consists of number of cancer patients those who registered themselves to the Web[5], with basic attributes such as sex, age, marital status, height and weight and an age group was taken from (15–65+) years. In this group of patients, major cancers were examined. By applying K-means [14, 15] and hierarchical agglomerative clustering (HAC) clustering algorithm [16], substantial distribution was found for incidence and mortality by sex and cancer site. The second clinical dataset consists of cancer patients that was taken from SEER datasets which has been recorded from the year 1975–2001 [17]. The data was classified into spatial dataset such as geographical, mobile phone usage data and non-spatial dataset (medical data). After this data mining algorithms were applied on the datasets like K-means, SOM and hierarchical clustering technique. The HAC performed better than other clustering methods regarding the results. The experimental results showed that data mining techniques are promising for clinical datasets clustering algorithms such as HAC and K-means in which HAC is applied on K-means are applied to determine the number of clusters. The quality of cluster is improved. Despite, the data analysis does not include missing records.

Santos et al. [18]

This study concentrates on a specific type of cancer who is liver cancer or (hepatocellular carcinoma HCC). A methodology is developed and presented to predict the

[5] www.dep.iarc.fr/globocan/database.htm.

one-year survival for patients with HCC. The work proposes a new cluster-based oversampling approach robust to database composed of heterogeneous clinical features with the presence of missing data.

The proposed methodology is based on data imputation process considering appropriate distance metrics for both heterogeneous and missing data by applying the heterogeneous Euclidean-overlap metric (HEOM) distance [19]. Then a clustering process is implemented in order to find naturally occurring clusters (or groups) within our HCC database using the K-means algorithm [14, 15] and the SMOTE algorithm [20] to build a representative dataset and use it as training example for different machine learning procedures (logistic regression LR and neural networks NN classifiers). The resulting approach is evaluated on dataset which comprises a set of N = 165 patients diagnosed with HCC with n = 49 features, selected according to the EASL-EORTC (European Association for the Study of the Liver—European Organization for Research and Treatment of Cancer, Clinical Practice Guidelines). The experiments were performed in order to show the feasibility of the proposed methodology to design survival prediction models for HCC disease and compare across baseline approaches that do not consider clustering and/or oversampling. The proposed methodology coupled with NN classifier presented better results than the other approaches cited in the article regarding all the performance measures (Accuracy, AUC and F-measure).

Yang and Chen [21]

In this study, data mining role in the diagnosis of lung cancer is revealed by taking the clinical information, which can be obtained, without surgery to replace the pathology report then use data mining techniques to find the correlation between the clinical information and the pathology report in order to support lung cancer pathologic staging diagnosis. As well as in the precedent work, a data imputation process is applied to delete the duplicates data and fill missing data points in the pathology reports and after the combination of both clinical and pathology report, a raw of dataset is comprised of clinical data and pathology report of lung cancer cases. The correlation between the clinical information and pathology reports is analyzed, and data association mining is employed to extract knowledge from the correlation between pathology reports and clinical information by using the a priori algorithm [22]. Many interesting rules are generated and evaluated using SCL and CLS approaches [23] and the more important that many interesting frequent patterns that make sense biologically are discovered through the experiment.

Shukla et al. [24]

This study discusses the use of data mining techniques on the diagnosis of breast cancer which is the most common cancer affecting females worldwide.

Breast cancer survivability prediction is of high priority around the world and a complex research task. Early diagnosis of cancer combined with early treatment makes better prognosis for cancer and identifying the status of this latter. According to this work, the availability of treatments and the high rate of incidents made survivability become a subject of much interest to health professionals and researchers [24].

This study contributes to cancer management as it engages knowledge discovery technologies to cancer patient records. Therefore, a robust data analytical model that can assist in a better understanding of breast cancer survivability in presence of missing data, providing better insights into factors associated with patient survivability, and establishing cohorts of patients that share similar properties is developed in this work. Unsupervised data mining methods are applied, such as the self-organizing map (SOM) algorithm [25] to consolidate patients into clusters of patients with similar properties. Next DBSCAN algorithm role came, to identify clusters among a set of records by detecting areas of high density [26]. Hence, a combination of two algorithm is used for segmenting patient records and created nine patient clusters with different survivability time which represent the basis for improving the survival prediction performances of multilayer perceptron (MLP) classifiers to be used for the training. As a result, a new data-driven approach is developed and the separation of patients into clusters improved the overall survival prediction accuracy based on MLP and revealed intricate conditions that affect the accuracy of a prediction.

Ma and Zhang [27]

Because of the heterogeneity and complexity of cancer patients in terms of disease progression, response to treatments, etc., defining subtypes of cancer and stratifying patient groups with the same disease but different subtypes for targeted treatments is of high importance for personalized and precision medicine and represents one major goal for cancer genomics. Unlike those using only one data type, approaches that incorporate multiomic data are more advantageous for patient clustering and disease subtype discovery despite their heterogeneity and noisy characteristic. For this, a new data clustering method is used in this study, and the issue related to heterogeneity of data is treated. An affinity network fusion (ANF) is presented to integrate multiomic data for patient clustering in this paper. ANF applies graph clustering to a constructed patient affinity/similarity matrix instead of patient-feature matrix and constructs an accurate" patient affinity network that incorporates information from multiple views. ANF is based on similarity network fusion (SNF) [28] who proved his efficiency in clustering patients but ANF is developed with more advantages than the latter. ANF generates better results with much less computation, and it provides also a more general framework for complex object clustering with multi-view data and can incorporate view weights. Furthermore, a semi-supervised model is developed and combined ANF and neural network (NN) for learning. In several cases, the model achieved very good results for few-shot learning (90% accuracy on test set with training less than 1% of the data). Due to this work, both unsupervised and semi-supervised affinity network fusion (ANF) framework that can integrate multiomic data for cancer patients clustering and subtype discovery are presented, and the potential for combining supervised and unsupervised learning through good representation learning is shown.

10.4 Others Studies and Approach's

In this section, we present the next five study based on other data mining technique which is supervised learning, despite the rareness of research work based on unsupervised learning, works related to supervising learning appears frequently in literature. Supervised learning methods are the act of training a model using observations on samples where the class of the outcome of interest is already known [8]. The resulting model is applied to new samples to predict values of the outcome for previously unseen observations and its performance evaluated by comparing predicted values to actual values for a set of test samples basing on specific metrics [8].

Mohammed et al. [29]

The present study treats one of the difficult to diagnose type of cancer which is nasopharyngeal carcinoma (NPC), a type of cancer in the head and neck and occurs in the esophagus section between pharynx and adenoidal opening and represent one of the important health problems in Malaysia. The diagnostic of NPC is a challenging issue that have not been optimally solved because of the complex structure of this type of cancer. The methodology and technology provided by data mining (DM) aim to transform the massive amount of data into useful information and new knowledge for decision making However, this work makes useful these techniques to develop the best solution and resolve the mentioned issue. The study presents an artificial neural network for automatic segmentation and identification of nasopharyngeal carcinoma from microscopy images without human intervention. This method uses, in the first stage, K-means after enhancing the image to be labeled in the regions based on their color. Secondly, neural network is used to select the right object on training stage. Finally, texture features for the segmented zone are extracted to the segmentation. To assess the efficiency of the anticipated results, an association of ANN and SVM segmentation results and automatic NPC classification is done. The evaluated of the framework is realized by comparing the automatic segmentation against the manual one, after this step the proposed segmentation solution is integrated into a classification framework for identifying benign and malignant tumor. As a result, the method is effective with high-level accuracy of classification of 91.01% and neural network classifier leads to significantly improved performance with the image texture features compared to the SVM classifier.

Wu et al. [30]

Another type of cancer is treated in this work, hepatocellular carcinoma (HCC) which is the sixth most common type of cancer and the third-leading cause of cancer-related deaths According to the statistics of the World Health Organization, the problem lies in the fact that HCC is frequently diagnosed in its late stages or cannot be treated through surgical resection. Therefore, radiofrequency ablation (RFA) has been proved to be an effective solution to treat patients with HCC recurrence, which cannot be treated through surgery. Here, it remains the importance of assessment patients after RFA. The aim of this study is to develop an artificial neural network

models with HCC-related variables to predict the one-year and second-year disease-free survival (DFS) of HCC patients receiving RFA treatments. A sample of patients who had received computer tomography-guided RFA is examined at National Taiwan University Hospital (NTUH), then 252 and 179 patients were divided into the one-year and second-year disease-free survival (DFS) groups with 15 categorical features were collected and used as inputs of ANN models. A multiple layer perceptron (MLP) model with the backpropagation learning rule was selected for model construction. As a result, the model is effective with acceptable prediction and high-level accuracy of 85.0% for one-year DFS prediction and with value of accuracy of 67.9% for two-year DFS prediction.

Iraji [31]

Lung cancer is the leading cause of cancer-related mortality worldwide.

Cardiothoracic surgery or thoracic surgery refers to operations on organs in the chest, like lungs and is frequently used to assess or repair lungs affected by cancer.[6] It is important for a patient with Lung cancer to know all possibilities of surviving after a thoracic surgery. In this stage, medical decision making is a serious challenge, about the risks of survival after the surgery. Therefore, this study aims to characterize and predict lung cancer survival after thoracic lung cancer. Many solutions to predict one year the postoperative survival expectancy in thoracic lung cancer surgery based on artificial intelligence are presented in this work. In this paper, researchers propose a method in the clinical diagnosis of thoracic lung cancer surgery, in order to help doctors in patient selection and identifies the risk of death in patients after surgery. For this purpose, A multi-layer architecture of sub-adaptive neuro fuzzy inference system (MLA-ANFIS) approach is implemented with various combinations of sixteen input features from the thoracic surgery dataset, neural networks, regression and ELM (extreme learning machine) (ELM base), ELM (RBF kernel), ELM (wave kernel), ELM (poly kernel) and ELM (lin kernel) based on the used dataset. The approach shows effective results to solve problem prediction of the postoperative survival expectancy in thoracic lung cancer surgery with accuracy of 88.78724% for ELM (wave kernel) comparing to other solutions cited in the article using the same dataset.

Guo et al. [32]

Early breast cancer recurrence detection improves a patient's chances of surviving, because this recurrence is a proof of poor response to adjuvant therapy and minimize those chances. The objective of this work is to establish a classifier that can predict whether a patient developed early breast cancer recurrence within 3 years of the initial diagnosis, based on similarities of their clinical and pathological diagnosis. The developed classifier is also able to group patients based on intrinsic properties of their diseases. According to their relevance to early relapse, for each subclass, the disease characteristics are listed in a hierarchal order. To build such classifier, a decision tree algorithm is applied on the clinical information dataset of a breast cancer patient containing pathological diagnosis, stroma percentage and expression

[6] www.cancercenter.com/treatments/thoracic-surgery.

of TGFß signaling components. Overall, a data mining approach is taken to generate a decision tree classifier to achieve the main goal wish which is to study whether stroma percentage and TGF signaling biomarkers are highly relevant to the status of early breast cancer relapse. Considering that these factors have different or even contrasting associations with breast cancer recurrence in subgroups of patients according to many studies cited in the article, decision tree classifier subdivided patients by hierarchically organizing breast cancer disease characteristics based on their relevance to early breast cancer relapse. The classifier achieves the aim and predicts correctly for 55 out of the 65 patients with 85% accuracy. The potential of the decision tree classifier as a powerful prognostic tool is proved in this work and its ability to be adopted in different academic and clinical settings

Tseng et al. [33]

This paper treats the use of data mining techniques in the case of ovarian cancer in Taiwan. According to this work, ovarian cancer is the second-leading cause of deaths among gynecologic cancers in the world. Furthermore, it is important for a patient to know how long he survives. Cancer is a preventable disease if it is analyzed at an early stage. Ovarian cancer has a 90% survival rate if its diagnoses take place in the first stage of the disease according to the article and that shows the importance of early detection to improve a woman's chances of surviving ovarian cancer. The recurrence of the cancer indicate that residual tumor cells are unresponsive to adjuvant therapy or have weak responses.

That recurrence is predicted with regard to several characteristics whose roles in ovarian cancer recurrence are controversial, and in the clinical symptoms of this devastating cancer, these allow for a better understanding of which variables are more potentially responsible to the relapse of ovarian cancer. The ultimate objective in this study was to identify and extract the risk factors resulting in the recurrence of the ovarian cancer, and then, the outcome is used as predictors for the five data mining classification techniques in this work.

The result is medical diagnostic model for diagnosing the recurrence of ovarian cancer. Five data mining classification techniques, including support vector machine (SVM), C5.0, extreme learning machine (ELM), multivariate adaptive regression splines (MARS) and random forest (RF), were used to achieve this goal. Each technique generates a sorted result of risk factors. The ranking results of five of the classifiers are combined to generate an overall ranking result. Eleven predictor variables served as the critical risk factors for constructing the E-C5.0, E-MARS, E-RF, E-SVM and E-ELM models. The E-C5.0 approach showed the highest average correct classification rate compared to other models with regard to classification accuracy of 90.00%. Must notice that the Chung Shan Medical University Hospital Tumor Registry was the source of medical records and pathology data collected in this study.

Different scalar metrics are computed to assess the performance of clinical diagnosis tests. Computing these metrics is based on several parameters such as **TP**, **TN**, **FP** and **FN**. A patient with (respectively without) the disease **D** predicted as such is designated as true positive (**TP**) (resp. true negatives (**TN**)). In case of wrong

predictions, the patients are false positives (**FP**) or false negatives, respectively (**FN**). Let **TP** (resp. **TN**) denote the number of true positives (resp. true negatives) and **FP** (resp. **FN**) the number of false positives (resp. false negatives). These quantities are usually illustrated in a matrix of confusion [34].

In this chapter, the considered metric to assess the performance of prognosis and diagnosis cancer tools is the Accuracy as mentioned in the Table 10.1. The accuracy (A) in defined as the rate of successful predictions and described by the formula (10.1) (Table 10.2)

$$A = \frac{TP + TN}{TP + FP + TN + FN}. \tag{10.1}$$

10.5 Conclusion

To face the tremendous flow of biomedical data produced from a day to another and make important and valuable the hidden knowledge within, the science of data management harnesses all their potentials and best efforts to develop methods and techniques for making use of these data and make from medical diagnostic reasoning a very important application area of intelligent systems. Cancer is one of the highest important biomedical domain research worldwide, and a wealth of works and studies applied to cancer are developed in literature makes the cited one in this chapter a drop in the sea. This chapter treats the importance of data mining to extract hidden knowledge from the amount of heterogeneous cancer datasets. Machine learning methods, techniques and tools are provided in this chapter and that can help solving diagnostic and prognostic cancer problems. The data mining played an important role in healthcare system of cancers disease such as prediction of disease progression, disease type or stage, extraction of hidden and useful knowledge, therapy planning and support and for the overall patient management. Moreover, DM techniques are used for cancer recurrence and mortality prediction as well as saving more lives by reducing death rates. Data mining in health care is used mainly in assisting for diagnosis for the doctors in making their clinical decision. The ethical issues presented in almost all works are related to missing values treatment and obtaining the highest accuracy among different used classifiers regarding the importance of this metric to evaluate the DM technique. Other challenge that arises in DM methods application in cancer domain is the control of overfitting and the scaling properties of these methods and if these latter can handle with large datasets and high-dimensional input (feature) and output (classes categories) spaces. Some solutions were proposed in works, which could resolve the mentioned issues.

To enhance the prediction and increase the accuracy, using hybrid models is proposed as solution: combining more than one data mining technique for diagnose and prediction or by training the system with new data from other different parts of the world. Developing prediction risk models must be an iterative process over time

Table 10.1 Overview of studies of the literature

References	DM model	MLT	Cancer type	Problem	Performance metric
Ma 2018 [27]	Descriptive	– Similarity network fusion		Integrate multiomic data for patient clustering into correct disease types.	
Shukla et al. 2017 [24]	Descriptive	– SOM algorithm – DBSCAN algorithm	Breast cancer	Survivability prediction for breast cancer patients	
Chang et al. 2017 [33]	Predictive	– Decision tree	Ovarian cancer	Identify risk factors and diagnose ovarian cancer recurrence	ACC = 90%
Guo et al. 2017 [32]	Predictive	– Decision tree	Breast cancer	Predict models for breast cancer recurrence	ACC = 85%
Mohammed et al. 2017 [29]	Predictive	– Artificial neural network	Nasopharyngeal carcinoma	Identify nasopharyngeal carcinoma	ACC = 91.01%
Iraji 2017 [31]	Predictive	– Artificial neural network	Lung cancer	Predict one year the postoperative survival expectancy	ACC = 88.87%
Chiu et al. 2017 [30]	Predictive	– Artificial neural network	Hepatocellular carcinoma	Predict the one-year and two-year disease-free survival (DFS) of HCC patients receiving RFA treatments	ACC = 85.0%
Santos et al. 2015 [35]	Descriptive	– k-means	Hepatocellular carcinoma	Improve survival prediction of hepatocellular carcinoma patients	

(continued)

Table 10.1 (continued)

References	DM model	MLT	Cancer type	Problem	Performance metric
Chen et al. 2015 [21]	Descriptive	– Association rules	Lung cancer	Identify correlation between clinical and pathology information for lung cancer diagnosis	
Chauhan et al. 2010 [13]	Descriptive	– k-means – HAC algorithm		Identify clusters in spatial cancer databases	

Table 10.2 Confusion matrix

Predicted as >	N	P
Negative (N)	TN	FP
Positive (P)	FN	TP

to adapt the system to new requirements and quality level. Some other challenges in DM area are related to data security, accessibility, incompleteness, inconsistent, noise and privacy concerns.

References

1. El Houby EMF (2018) A survey on applying machine learning techniques for management of diseases. J Appl Biomed
2. Schutte AE (2017) Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the global burden of disease study
3. Murdoch TB, Detsky AS (2013) The inevitable application of big data to health care. JAMA 309(13)
4. Brin S (1998) Extracting patterns and relations from the world wide web paper presented at the international workshop on the world wide web and databases. Springer, Berlin
5. Huang M, Zhu X, Ding S, Yu H, Li M (2006) ONBRIRES: ontology-based biological relation extraction system. In: Proceedings of the APBC
6. Roser M, Ritchie H (2018) Cancer. Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/cancer. (Online Resource)
7. Vidyasagar M (2017) Machine learning methods in computational cancer biology. Ann Rev Control 43:107–127
8. Callahan A, Shah NH (2017) Machine learning in healthcare. Key Adv Clin Inform 279–291
9. Magoulas GD, Prentza A (2001) Machine learning in medical applications. Lecture Notes in Computer Science, 300–307
10. Jothi N, Rashid NA, Husain W (2015) Data mining in healthcare—a review. In: Proceedings of the 3rd information systems international conference, pp 306–313
11. Patel S, Patel H (2016) Survey of data mining techniques used in healthcare domain. Int J Inf Sci Tech 6:53–60

12. Kantardzic M (2011) Data mining: concepts, models, methods, and algorithms, 2nd ed. Wiley-IEEE Press
13. Chauhan R, Kaur H, Alam MA (2010) Data clustering method for discovering clusters in spatial cancer databases. Int J Comp Appl 10(6):0975–8887
14. Kaufinan L, Rousseeuw PJ (1990) Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, Hoboken
15. Chen MS, Han J, Yu PS (1996) Data mining: an overview from database perspective. IEEE Trans Knowl Data Eng 5(1):866–883
16. Zhao Y, Karypis G (2002) Evaluation of hierarchical clustering algorithms for document datasets. In: Proceedings of the eleventh international conference on information and knowledge management, McLean, Virginia, pp 515–524
17. Surveillance, Epidemiology, and End Results (SEER) program (www.seer.cancer.gov) public-use data (1973–2002), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2005
18. Santos MS, Abreu PH, García-Laencina PJ, Simão A, Carvalho A (2015) A new cluster-based oversampling method for improving survival prediction of hepatocellular carcinoma patients. J Biomed Inform 58:49–59
19. Batista GE, Monard MC (2003) An analysis of four missing data treatment methods for supervised learning. Appl Artif Intell 17:519–533
20. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority oversampling technique. J Artif Intell Res 16(1):321–357
21. Yang H, Chen Y-PP (2015) Data mining in lung cancer pathologic staging diagnosis: correlation between clinical and pathology information. Expert Syst Appl 42(15–16):6168–6176
22. Chai S, Yang J, Cheng Y (2007) The research of improved Apriori algorithm for mining association rules. In: Paper presented in the IEEE international conference 540 on service systems and service management, pp 1–4
23. Hipp J, Güntzer U, Nakhaeizadeh G (2000) Algorithms for association rule mining—a general survey and comparison. ACM SIGKDD Explor Newsl 2(1):58–64
24. Shukla N, Hagenbuchner M, Win KT, Yang J (2018) Breast cancer data analysis for survivability studies and prediction. Comput Methods Programs Biomed 155:199–208
25. Self-Organizing Maps (2001) Springer-Verlag, New York, Inc.
26. Ester M, Kriegel H, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceeding of the 2nd international conference on knowledge discovery and data mining (KDD'96), pp 226–231
27. Ma T, Zhang A (2018) Affinity network fusion and semi-supervised learning for cancer patient clustering. Methods 145:16–24
28. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, Haibe-Kains B, Goldenberg A (2014) Similarity network fusion for aggregating data types on a genomic scale. Nat Methods 11(3):333–337
29. Mohammed MA, Ghani MKA, Hamed RI, Ibrahim DA, Abdullah MK (2017) Artificial neural networks for automatic segmentation and identification of nasopharyngeal carcinoma. J Comput Sci 21:263–274
30. Wu CF, Wu YJ, Liang PC, Wu CH, Peng SF, Chiu HW (2017) Disease free survival assessment by artificial neural networks for hepatocellular carcinoma patients after radiofrequency ablation. J Formos Med Assoc 116:765–773
31. Iraji MS (2017) Prediction of post-operative survival expectancy in thoracic lung cancer surgery with soft computing. J Appl Biomed 15(2):151–159
32. Guo J, Fung BCM, Iqbal F, Kuppen PJK, Tollenaar RAEM, Mesker WE, Lebrun J-J (2017) Revealing determinant factors for early breast cancer recurrence by decision tree. Inf Syst Front 19(6):1233–1241
33. Tseng C-J, Lu C-J, Chang C-C, Chen G-D, Cheewakriangkrai C (2017) Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence. Artif Intell Med 78:47–54

34. Witten IH, Frank E, Hall MA, Pal CJ (2005) Data mining: practical machine learning tools and techniques
35. Santos MS, et al. (2015) A new cluster-based oversampling method for improving survival prediction of hepatocellular carcinoma patients. J Biomed Inf 58:49–59A