

Mobile services infrastructure for frailty diagnosis support based on Gower's similarity coefficient and treemaps

Jesús Fontecha*, Ramón Hervás and José Bravo
MAMI Research Lab, Castilla-La Mancha University, Ciudad Real, Spain

Abstract. The early detection and diagnosis of frailty is important in improving the quality of elderly life. Physicians must take many factors into account to achieve an objective and standardised method of frailty assessment. In this work, we present a novel mobile service infrastructure to improve the accuracy of frailty diagnosis in a population of elders, focusing on the main frailty risk factors and applying similarity algorithms based on Gower's coefficient. In addition, we have developed a mobile application to obtain the assessment results from the implemented services. The results are displayed on the mobile device using hierarchical visualisation structures known as treemaps. The system has been evaluated as a frailty diagnosis tool in a group of elderly patients.

Keywords: Frailty, elderly people, mobile service, similarity coefficient, treemap visualisation

1. Introduction

Frailty in the elderly is closely related to aging and dependence. Aging leads to frailty, which may result in a state of dependence. The natural process of aging is characterised by a loss of physical, cognitive and other abilities. However, an objective definition of frailty remains elusive. In 1988, Woodhouse [25] defined the frail elderly as individuals over 65 years of age who depend on others for their basic needs. Gillick [7] defined frail elderly people as “old debilitated individuals who cannot survive without substantial help from others”, emphasising the social consequences of frailty. Frailty is a condition that increases the risk of disability and dependency in the elderly. In contrast to old age, frailty and dependence are often reversible states; however, aging is the main cause of frailty. Fried [5] proposes a phenotype of frailty based on the clinical observed symptoms and signs of the syndrome. Fried specifies out five general criteria to determine whether an elder is frail: weight loss, weakness and fatigue, decreased grip strength, slow gait speed and low physical activity.

Many variables must be considered in the detection and diagnosis of frailty, including the following factors:

- Medical: Presence of chronic diseases, gait disturbance, sensory deprivation, recurrent falls, inaccurate perception of health, polypharmacy, hospitalisation.

*Corresponding author: Jesús Fontecha, MAMI Research Lab, Esc. Sup. de Informática, Paseo de la Universidad, 4 13071 Ciudad Real, Spain. E-mail: jesus.fontecha@uclm.es.

- Functional: Dependence in the basic activities of daily living (BADL) or activities of daily living (ADL) and instrumental activities of daily living (IADL).
- Socio-economic: Living alone, recent widowhood, age over 80 years, low incomes.
- Cognitive: Depression, cognitive impairment.
- Institutional: Admission to elderly homes.

This list includes a large number of parameters and variables that must be considered, as discussed extensively by Hamerman [12]. Frailty is currently studied based on some but not all of the above criteria. Obtaining a global and accurate index or indicator to define the state of frailty in an elder is a complex task based on the following information:

- Global scores from standard questionnaires completed by doctors
- An overview of the elder and his environment
- Measures obtained using medical instruments
- Analysis of lab reports from the elderly patient

In practice, doctors use only a small portion of the parameters related to these items in the frailty assessment, and use their subjective experience as the main tool in the final diagnosis.

The use of mobile technologies and service-oriented approaches provide new methods for diagnosis support in frailty assessment and other medical areas. We therefore adopted the following hypothesis in this research: essential information regarding the frailty assessments of individuals in an elderly population can be obtained by objectively analysing a broad set of relevant parameters through different algorithms implemented as web services and deployed in mobile devices. In this paper, we present a novel mobile service designed to generate accurate and centralized frailty assessments. We focus on the most common frailty parameters according to the literature and the geriatricians experience, as described in the following sections, and implement our proposal using similarity algorithms. The final results of the frailty assessment are displayed on the mobile device using treemaps.

The paper is organised into 7 sections. Section 2 discusses previous work on the use of new technologies in healthcare and existing rudimentary approaches to frailty diagnosis, including classification methods based on the similarity concept. In Section 3, we provide an overview of the cluster analysis stages and similarity algorithms. We describe the application of Gower's similarity algorithm to frailty assessment in Section 4. In Section 5, we describe the proposed service in detail, including its deployment on a mobile device. Section 6 evaluates the proposal based on experimental results from a group of selected elders. Finally, we conclude in Section 7.

2. Related work

New technologies are currently used in healthcare environments for many purposes, including telemedicine, patient monitoring, location and identification tasks, emergencies and clinical information management among others. The incorporation of mobile devices is revolutionising this heterogeneous field, and there is a growing body of research on the use of mobile technologies in supporting the daily medical tasks of individuals. For example, Villarreal [24] proposes a mobile system for patient diabetes monitoring. Pinnock [17] introduces an asthma monitoring system using mobile phones, which enables the patient to self-manage his or her disease through feedback and reminders of appropriate actions. Similar systems have been proposed for the mobile monitoring of many types of diseases and clinical tasks. However, the use of new technologies and mobile devices is not widespread in frailty domain.

In the study of frailty, it is important to identify the relevant factors in the frailty assessment. In Section 3.1, we describe the factors used in our work, based on the literature and the variables most commonly included in related research. Several alternative approaches to frailty analysis have also been proposed. Jones [14] introduces a frailty index based on a detailed geriatric assessment, in which a set of variables such as balance, communication, cognitive state, nutrition, continence, ADLs and comorbidity are co-analysed. However, Jones concluded that the optimal approach to frailty assessment remains undetermined. Along similar lines, Searle et al. [18] proposed a frailty quantification procedure based on a dataset including coded non-numerical variables.

Gobbens et al. [8] introduce a conceptual framework in which the most important factors related to frailty are determined experimentally. The experiments measure the cognitive function, strength, balance, nutrition, physical activity and mobility of the patient; social and psychological factors are emphasised less in this study. The importance of the physical condition of the patient in frailty assessments is emphasised in many studies including variables measured while the patient is engaged in physical activity. Accelerometer-based applications for the collection and analysis of movement data are therefore an important component of mobile devices designed to assess frailty (see Section 3.1).

Following the collection of the data on the frailty variables, a thorough analysis must be performed to provide useful support to the end users, in this case physicians. Many methods are available for analysing the entire set of selected variables. Two of the main techniques based on observations are cluster analysis and classification mechanisms. Gower's method for assessing the similarity between each pair of items in a population is one of the most powerful analysis methods owing to its favourable properties [15]. For example, Gonçalves et al. [9] use the Gower distance (in addition to several other measures) to determine the genetic similarity between several tomato seeds.

Most studies that use similarity methods are focused on the analysis in characteristics between different species of animals or plants. However, these methods can also be extended to other areas of research. In a clinical context, classification and similarity mechanisms are used in the analysis of diseases. For example, Socransky [20] presents an approach based on cluster analysis and the Gower coefficient to estimate the extent of periodontal disease in 22 patients. Gelnarová [6] performs a comparison between three classification methods, including one of which uses the Gower coefficient of similarity, in the diagnosis of very early stage prostate cancer in a group of 826 patients. In these examples, similarity mechanisms based on Gower's coefficient are applied in a clinical context. However, frailty analysis using classification methods is not widespread. To our knowledge, classification methods and similarity concepts have not been used previously in mobile technologies for decision making related to frailty diagnosis.

3. Cluster analysis and similarity algorithms

Clustering is a classification technique that divides data into groups, known as clusters, which are meaningful, useful, or both. These groups of objects share common characteristics and play an important role in how people analyse and describe the world. Human beings are skilled at dividing objects into groups (clustering) and assigning particular objects to these groups (classification) [22]. In automatic clustering, each group must be homogeneous with respect to several characteristics, i.e., the observations within each group must be similar to one another.

Cluster analysis algorithms involve three major stages:

1. Selection of the relevant variables. This stage allows us to characterise the objects that will be categorised based on relevant criteria.

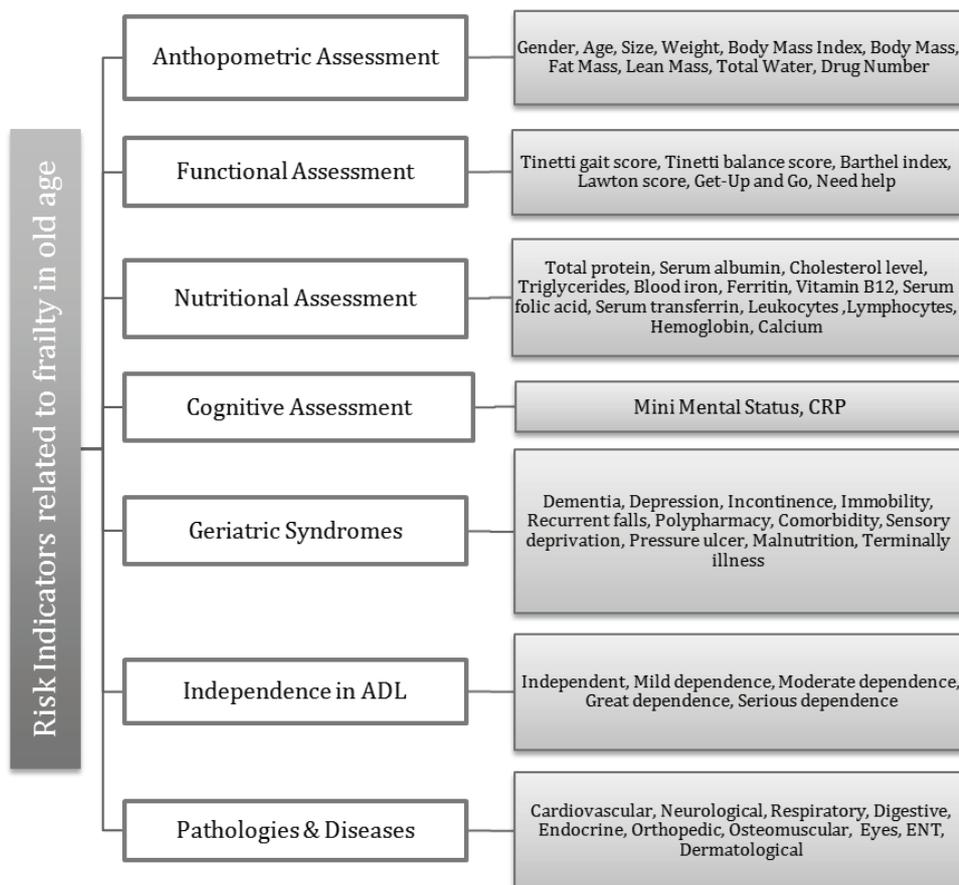


Fig. 1. Relevant clinical variables in frailty assessment. ADL indicates: Activities of daily living; CRP indicates: Cognitive scale from “Cruz Roja Spain” at <http://www.cruzroja.es>; ENT indicates: Ear, Nose and Throat disorders.

2. Normalisation of the variables. A normalisation procedure is often necessary because the selected variables may have different types (e.g., quantitative vs. qualitative) and units.
3. Calculation of the similarity measures. The similarity measure indicates the strength of the relationship between two objects.

We consider a single initial group, described in detailed in Section 6, to illustrate the function of the developed service.

3.1. Selection of the relevant variables

In the frailty diagnosis literature, certain factors are more important than others. Espinoza [2], for example, identifies a group of possible risk factors based on the frailty phenotype and a systematic review of previous works. However, the relative importance of the various factors is not discussed in a quantitative fashion.

On the one hand, most variables related to frailty are obtained from the patient record (e.g., from test scores and scale ratings, results of lab reports, and general patient data). Social and psychological factors are not typically considered because they are not a directly connected to the patient record. Many doctors

Table 1
Classification of frailty factors

Binary (values)	Qualitative (values)	Quantitative – numeric value
Requires Help (No/Yes), Gender (Male/Female), Dementia (0/1), Depression (0/1), Incontinence (0/1), Immobility (0/1), Recurrent falls (0/1), Polypharmacy (0/1), Comorbidity (0/1), Sensory deprivation (0/1), Pressure ulcer (0/1), Malnutrition (0/1), Terminal illness (0/1), Cardiovascular problems (0/1), Neurological problems (0/1), Respiratory problems (0/1), Digestive problems (0/1), Endocrine (0/1), Orthopaedic problems (0/1), Osteomuscular problems (0/1), Eye disease (0/1), ENT (0/1), Dermatological problems (0/1)	Independence in ADL (independent/mild dependence/moderate dependence/major dependence/serious dependence)	Age, Size, Weight, BM, BMI, Fat mass, Lean mass, Total water, Drug number, Tinetti gait score, Tinetti balance score, Barthel index, Lawton score, Get-up and go, Total protein, Serum albumin, Cholesterol, Triglycerides, Blood iron, Ferritin, Vitamin B12, Serum folic acid, Serum transferrin, Leukocytes, Lymphocytes, Hemoglobin, Calcium, MMS, CRP

and geriatricians agree that frailty assessment must incorporate variables from several clinical areas. Figure 1 shows the most common areas and variables taken into account in clinical frailty assessment.

The value of each factor may be binary, discrete or continuous. In Table 1, we classify each variable as binary, qualitative or quantitative according to the geriatric criteria. Thus, for example all geriatric syndromes are considered binary because they may or may not appear. There are cases such as “Respiratory problems” in which a single symptom determines the true value (or 1) and other such as “immobility” evaluate that the patient can not move without help (e.g., if he is bedridden). The type of each variable must be provided as input to the similarity algorithm.

Moreover, not all variables have the same degree of importance in frailty assessment. Geriatricians typically evaluate the importance of each variable depending on the context and particular case. For example, a higher emphasis may be placed on functional variables in a case where the patient is limited by functional deficits. Similarity algorithms allow different weights to be assigned to different variables.

According to most practitioners, physical activity analysis is essential because it offers substantial information regarding the functional level of an individual. In particular, gait analysis using the tests introduced by Tinetti [23] and Barthel [1] provides a reliable measure of frailty. The development of accelerometers and their integration into mobile devices therefore facilitates an objective analysis of the physical conditions of a broad range of elderly individuals. Accelerometer sensors certainly provide far more information compared to current tests. In [4], we describe a set of characteristics known as *dispersion measures* that can be determined from the values provided by the accelerometer. These new quantitative variables (each calculated for the three coordinate axes) are as follows: *Arithmetic mean*, *Standard deviation*, *Amplitude*, *Absolute mean difference*, *Acceleration mean*, *Variance* and *Pearson’s coefficient of variation*.

All of these variables will be included in the new *dispersion measures* group of risk indicators for frailty assessment.

3.2. Normalisation of the variables

All of the identified variables are normalised or standardised before the similarity calculation so that different instances of a variable (e.g., weight) are measured in the same units (e.g., kilos). Furthermore, the formulas used to determine the similarity coefficients are different depending on the type of each variable (binary, qualitative or quantitative), as discussed in the next section. Therefore, if the patient records included variables measured using different metrics, then the system requires a normalisation procedure to convert all of the measurements to the same units.

3.3. Calculating the similarity measures

One of the main components of clustering analysis is the calculation of the degree of similarity between individuals. There are several methods of calculating the similarity matrices, dissimilarities and distances [3] between individuals in a population.

The appropriate similarity algorithm may be different in each case, depending on the type of the variables (binary, qualitative or quantitative). For example, if the entire set of variables is binary, then the Jaccard similarity coefficient [13] is one of the most useful measures. If all of the variables are qualitative, then we must choose an alternative method. In this case, Everitt [3], for example, recommends an explicit conversion of the previous variables to binary form. However, when mixed variables (binaries, qualitative and quantitative) are included Gordon [10] for example, suggests converting the different variables to the same scale (binary). In addition, there is general agreement on the use variables of different types only as a complement to interpret the final results [21].

Gower's general similarity coefficient is currently one of the most popular measures of proximity or similarity for mixed data types. The Gower coefficient [11] make it possible to determine the degree of similarity between two individuals or cases (i,j) represented by binary, qualitative and/or quantitative data. The Gower coefficient is defined in Eq. (1).

$$S_{ij} = \frac{\sum_{k=1}^p W_{ijk} S_{ijk}}{\sum_{k=1}^p W_{ijk}} \quad (1)$$

where:

- S_{ijk} denotes the contribution from the k th variable.
- W_{ijk} is usually 1 or 0 depending on whether the comparison is valid for the k th variable. If differential variable weights are specified, then W_{ijk} is the weight of the k th variable. If the comparison is not valid, then W_{ijk} is 0.
- The effect of the denominator is to divide the sum of the similarity scores by the number of variables (or by the sum of their weights if the variable weights are specified).

Gower defines the value of S_{ijk} for quantitative variables as in Eq. (2).

$$S_{ijk} = 1 - \frac{|x_{ik} - x_{jk}|}{R_k} \quad (2)$$

where:

- R_k is the range of values for the k th variable. For quantitative variables S_{ijk} ranges between 0 and 1, applying the formula $|x_{max} - x_{min}|$.

For binary and qualitative data, we take $S_{ijk} = 1$, if $x_{ik} = x_{jk}$, or $S_{ijk} = 0$, if $x_{ik} \neq x_{jk}$.

As noted above, the weight W_{ijk} for the comparison of the k th variable is usually 1 or 0. However, if we assign differential weights to the variables, then W_{ijk} is either the weight of the k th variable or 0, depending on whether the comparison is valid. The similarity calculation can therefore account for the relative importance of each variable (given by the weight).

If the weight of any variable is zero, then the variable is effectively ignored in the calculation of the proximities and similarities. These variables are “masked” in the clustering and classification; however, they may be taken into account to assist in the interpretation of the cluster analysis.

Given the similarity coefficient, we can also determine the dissimilarity coefficient between two individuals as in Eq. (3):

$$D_{ijk} = 1 - S_{ijk} \quad (3)$$

Table 2

An excerpt of frailty factors for 4 patients. The factors considered in this example are the following: Weight, Body Mass Index (BMI), level of Activities of Daily Living (ADL), Depression syndrome, and Total Protein and Blood Iron as biological indicators

	Weight (Kg)	BMI	ADL level	Depression	Total protein (g/dl)	Blood iron
Patient 1	74,5	26	Independent	0	7,2	125
Patient 2	67	28	Independent	1	7,1	59
Patient 3	54,8	29	Independent	0	6,4	104
Patient 4	57,3	22	Mild dependence	0	6,4	32

$$S_{12} = 1 * (1 - \frac{|74,5 - 67|}{19,7}) + 1 * (1 - \frac{2}{7}) + 0,8 * (1 - \frac{0,1}{0,8}) + 0,75 * (1 - \frac{66}{93}) + (0,1 * 1) + (0,1 * 0) = 0,63$$

[value1 - value2] Rweight weights Sum

Fig. 2. Gower's coefficient formula for similarity calculation based on data from Table 2.

4. Application of Gower's similarity coefficient

Gower's algorithm comprises the core of the service developed in this paper. The calculation of the similarity coefficients from the patient record data involves working with mixed variables and obtaining diagnostics depending on the particular situation. Gower's algorithm is therefore suitable in this study. The properties of Gower's Similarity Coefficient have been extended as follows by Londoño [16]:

- Most of the clinical characteristics of the species involve variables of multiple types (quantitative, qualitative, binary), and Gower's Coefficient is suitable for calculating similarities for mixed variables.
- Gower's Coefficient is appropriate for databases in which there are missing values or observations. These missing values do not "break" the calculation of the final results. This property is very useful in taxonomic studies due to the frequent occurrence of missing observations or values.
- Through the use of this coefficient, it is possible to weigh the variables independently, depending on the importance assigned to each variable at each moment. Physicians can thereby perform a frailty assessment focusing on specific areas such as physical, nutritional, cognitive or anthropometric factors.

Based on the properties of Gower's coefficient, we propose that the similarity coefficient be calculated accounting for the frailty variables that were defined previously in Section 3.1. An example similarity calculation using Gower's Algorithm with a limited number of variables (see Table 2), is shown in Fig. 2 for 2 patients.

In this case, there are four quantitative variables (weight, body mass index (BMI), total protein, blood iron), one qualitative variable (ADL level) and one binary variable (depression). the weight and BMI variables have an importance of 100% (their weights are 1), while the weights of the remainder of the variables are as follows: 0,8 (importance of 80%) for total protein, 0,75 for blood iron, and 0,1 for the



Fig. 3. Frailty mobile application screenshots. (a) Movement analysis and calculation of the dispersion measures related to gait exercise using the accelerometer built into the mobile phone. (b) Visualisation of the frailty variable values. High values are coloured in red, low values are coloured in blue, variables without values are coloured in grey and, variables in green represent normal values. (c) Global weights of each group of variables. (d) Editing of specific weights for the anthropometric variables group.

ADL level and depression variables. Based on Gower's formula, the similarity result between patient 1 and patient 2 is 0,63. Using this algorithm for the entire group of patients, we can obtain a similarity matrix based on Gower's coefficient as in Eq. (4). Geriatricians are thereby provided with a general method for determining the frailty state of an elderly patient in comparison to other patients in an elderly population.

$$S = \begin{bmatrix} 1 & 0,63 & 0,36 & 0,18 \\ 0,63 & 1 & 0,49 & 0,34 \\ 0,36 & 0,49 & 1 & 0,52 \\ 0,18 & 0,34 & 0,52 & 1 \end{bmatrix} \quad (4)$$

Clearly, each patient can be studied several times, supplying new values for the frailty variables. We therefore use an instance database rather than a patient database, where an instance consists of a complete set of frailty variables associated with a given patient at a given moment [4].

5. Service implementation and deployment

The main goal of our service is to obtain a group of similarity results embedded in representable structures on a mobile device. We therefore propose the initial development of a mobile application to centralize the tasks involved in frailty assessment. The application must collect the relevant variables from the patient record, facilitate the visualisation of the variables, configure the inputs and weight values for a specific frailty analysis, and calculate the new group of relevant variables related to dispersion measures (see Section 3.1) using an accelerometer-enabled mobile phone. All of these features have been

implemented in an Android¹ mobile application. Figure 3 shows several screenshots of this application for a particular patient instance.

Using a service-oriented approach, the information collected by the mobile phone is sent to the web-service back-end to be analysed, calculating the corresponding frailty results and sending these ones to the mobile application at run time through the network. We have developed the web services required for this fact as software components of a more complete system. In this section, we will focus on the Gower's similarity service and the visualisation of the final results on the mobile device. The Gower similarity service is divided into five stages, as described below.

5.1. Ranges calculation

First, we must determine the range of each quantitative variable. Given a group of variables, when every variable in the group is quantitative, we can apply Eq. (2), in Section 3.3. However, there are both groups whose variables are quantitative and qualitative (such as the functional group) and groups in which all of the variables are qualitative or binary (the geriatric syndromes and ADL level groups). In the case of groups with mixed variables, we discriminate qualitative and binary variables. Algorithm 1 shows the pseudocode for a typical range calculation on groups with mixed (or only quantitative) variables. The output of the algorithm consists of an array of the calculated ranges for each variable in each group.

Algorithm 1 Range Calculation Algorithm (for a specific group of variables)

```

Define double Array of ranges
Connect to the System Database
Get frailty variables of the group (for all stored instances)
for each variable do
  if variable is quantitative then
    Get Range  $|MAX(x_i) - MIN(x_i)|$ 
    Cast to double type (if necessary)
    Save into double Array of ranges
  end if
end for
Return double Array of ranges.

```

In the case of the *dispersion measures* group, we must also consider the specific activity (e.g., gait) to be an input argument in the previous algorithm because the values of each dispersion variable may refer to more than one stored activity.

5.2. Weight calculations

First, the weights of all variables are initialised to one by default, i.e., the maximum of importance (100%) is assigned to all of the variables. However, changes in the weights may be required by the physician following a more thorough analysis in a particular domain (e.g., the anthropometrical domain). In this case, each individual variable may have a specific weight. When we perform a study of a patient, we account for each specific instance from that patient through the weights assigned to each group of variables (the total weight of a group is given by the sum of the weights of all variables in that group).

¹<http://www.android.com/>.

Clearly, the assigned weights depend on the doctor. We therefore propose the development of programming mechanisms that are accessible from the mobile application user interface (UI) to facilitate this task, as in Figs 3b and 3d.

Algorithm 2 shows the pseudocode for the calculation of the weights of the frailty variables for a specific instance and a given group.

Algorithm 2 Weight Calculation Algorithm (for a specific group of variables and a given instance)

Define *Weight Array* (where a weight object consists of a variable name and percentage value).

Connect to the System Database

Get the weights of the frailty variables of the group (for the specified instance)

for each variable **do**

if no weights assigned, **then**

 Create and store a new weight with *importance* = 1 by default

end if

 Get the name and weight of the variable to create a weight object

 Save in *Weight Array*

end for

Return *Weight Array*

Note that the sum of all variable weights is a required input in Gower's formula. Following the weight calculation, it is therefore necessary to obtain the sum of the weights for each group of variables.

5.3. Calculation of Gower's coefficient

The next step in the proposed service is the calculation of Gower's Coefficient. Based on the assumptions described in Section 4, the calculation of the global value of Gower's Coefficient requires the computation of the partial coefficient numerators for each group of frailty variables. Algorithm 5.3 shows the pseudocode for the global Gower's Coefficient calculation. This calculation is always performed on the patient instance under study in relation to a second instance.

Algorithm 3 Global calculation of Gower's Coefficient

Define *double Gower*

Define *double Sum*

Define *double WeightsSum* = Sum of variable weights

for each group of frailty variables **do**

 Sum = Sum + Gower's Coefficient Numerator

end for

Gower = Sum / *WeightsSum*

Return *Gower*

An example of the partial coefficient calculation (corresponding to a partial sum of the numerator values, underlined in Algorithm 3) is shown in Algorithm 4, for a mixed group (including both quantitative and qualitative variables). In this case, each variable is related to a *measure object* (in the object-oriented programming – OOP paradigm) with the following attributes: name, value, minimum reference value, maximum reference value and units. This partial calculation is performed for two instances using the *Weight Array* from the first instance (or target instance).

The part of the service based on the above algorithms provides a single value for Gower's Coefficient between two instances, accounting for every frailty variable included in the system (that has a value). However, we must determine all of the similarity coefficients between the current instance and the remaining instances. In the following sections, the previous calculations are generalised, and the nodes calculation task is performed based on the similarity results.

Algorithm 4 Partial calculation of the numerator of Gower's coefficient

```

Define double PGowerNumerator
Get Measures Array1 from the first instance
Get Measures Array2 from the second instance
for each weight of Weight Array do
  if measure associated with that weight is qualitative/binary then
    if measure.value from Measures Array1 = measure.value from Measures Array2 then
       $PGowerNumerator = PGowerNumerator + (weight * 1)$ 
    else
       $PGowerNumerator = PGowerNumerator + (weight * 0)$ 
    end if
  else
    value1 = measure.value from Measures Array1
    value2 = measure.value from Measures Array2
    range = Get range for measure
    if none of the ranges are zero, then
       $PGowerNumerator = PGowerNumerator + (weight * (1 - |value1 - value2|/range))$ 
    end if
  end if
end if
Return PGowerNumerator

```

5.4. Node calculations

The node calculation task uses a new type of element known as *node*, which can be encapsulated in a programming object. We propose to create a tree structure consisting of nodes to store the results in an organised fashion, where the root node refers to the patient instance to be evaluated.

A node object contains the following attributes:

- Parent Instance Id., corresponding to the instance identifier of the parent node.
- Instance Id., an identifier of the current instance associated with the node.
- Age, referring to the age of the patient represented by the node.
- Similarity Coefficient, the value of Gower's coefficient between the parent node and the current node.

Algorithm 5 shows the pseudocode for the calculation of the nodes for a specific initial instance in relation to the remaining instances. This algorithm returns a list of nodes that are ordered by their similarity coefficients (in decreasing order).

Algorithm 5 Calculation of the nodes

```

Define Node List
Define initial Instance
Get stored instances
for each instance that is not equal to the initial instance do
  Calculate Gower's Coefficient
  Get Node Age
  Create the new Node
  Add the new Node to Node List
end for
Sort Node List (desc.)
Return Node List

```

Following the calculation of the nodes, it is possible to create a tree of nodes (ordered by the similarity coefficient) by iterating with each successive node in the list being taken as a new parent node. Clearly,

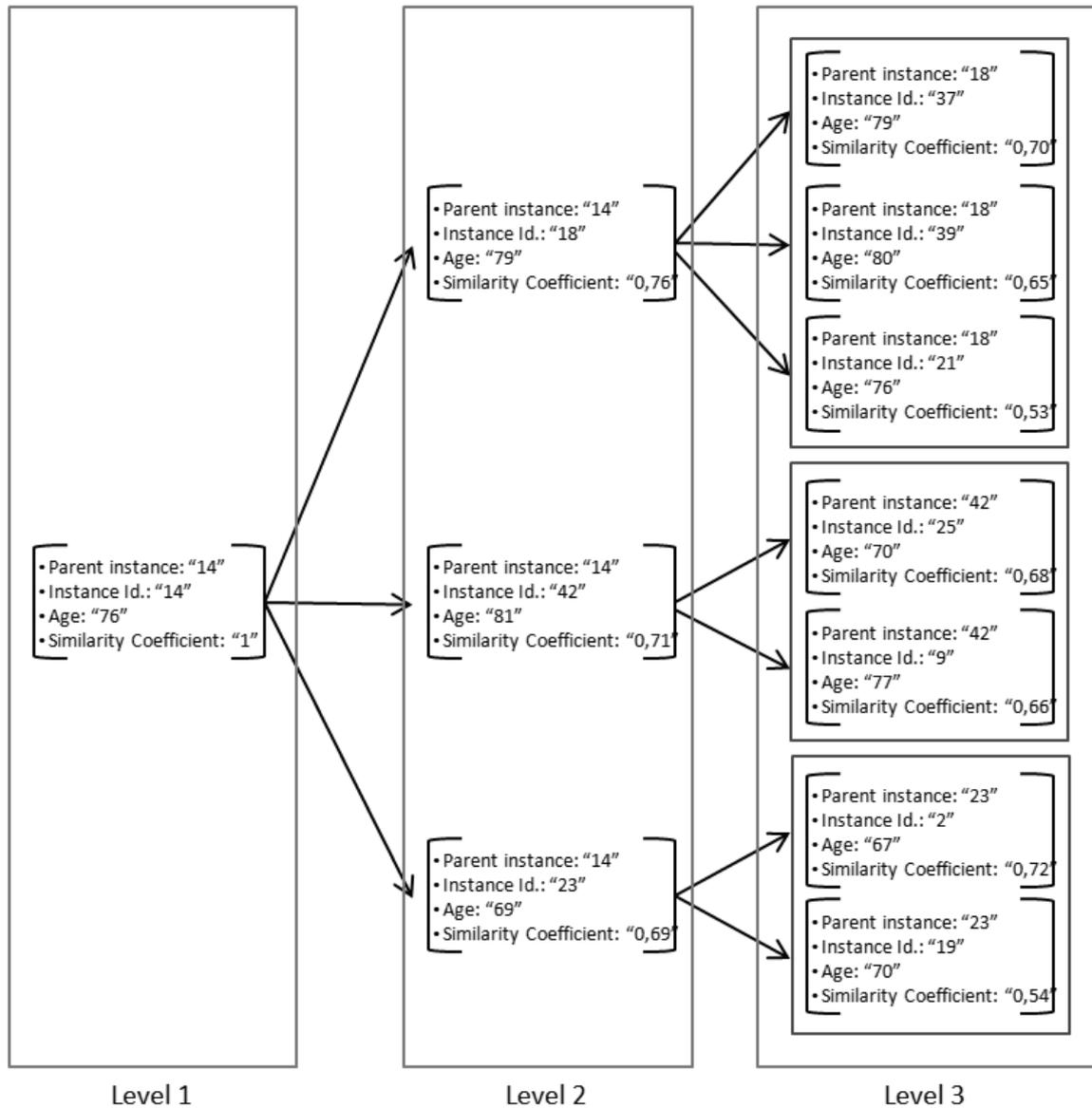


Fig. 4. Example of the final tree structure with $S_G \geq 0,52$ and 3 depth levels. Each group of child nodes is ordered by the value of the similarity coefficient.

when the number of instances is large, the system will require more processing capacity, more storage resources, better memory management and more time to generate the similarity results. In addition to the optimisation of these system features, we propose to constrain the depth of the tree to a maximum of three levels and restrict the number of child nodes (for each parent node) to three. With a higher number of tree levels and child nodes, the proposed service may require an unfeasible amount of processing time (depending on the stored instances), and the final results may be less useful to the physician because the similarity coefficient values are progressively lower (in relation to the instance under study, corresponding to the root node) at each tree level.

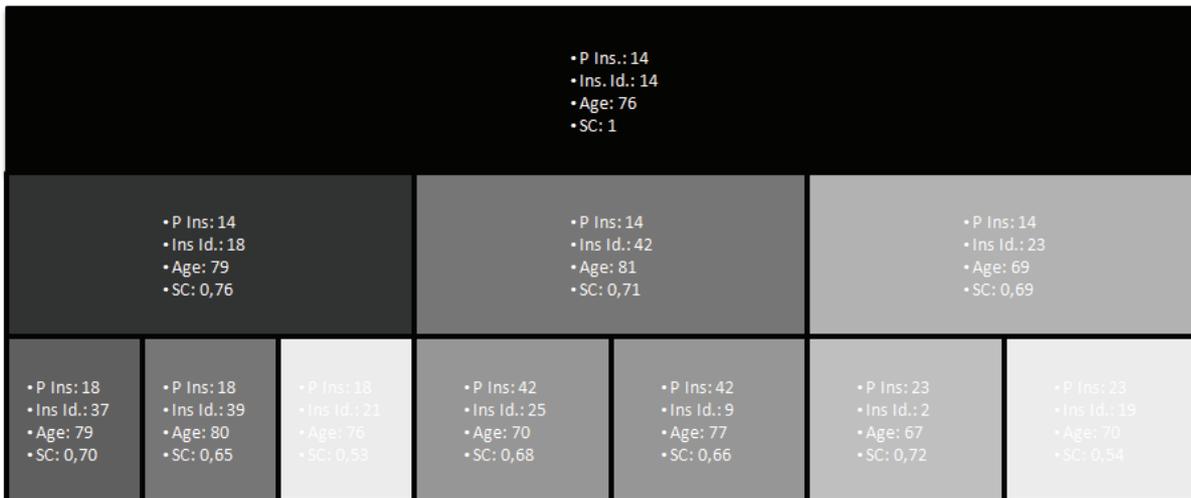


Fig. 5. Example of the treemap view corresponding to a result tree from the node list.

The tree depth (with values ranging from one to three) can be set up from the user interface of the mobile frailty application by the doctor, as for the weights of the frailty variables. The instances whose similarity values are lower than a given value can also be ignored in the construction of the tree using the mobile application. Figure 4 shows a graphical example of a tree provided by the implemented service, using all similarity coefficients exceeding 0,51. This value has been selected as the minimum similarity value to illustrate the behaviour of the service when there are nodes with similarity values higher and lower than the threshold.

5.5. Treemap visualisation

The mobile device receives the previous node list, represented by a tree structure resulting from the frailty assessment based on the analysis of a specific instance. In the Treemap approach [19], we can optimise and compress the visualisation of the results for mobile screens using treemaps. A hierarchical structure is thereby constructed from the node tree, and the results are displayed to the end user (physician) in a more friendly and intuitive fashion.

Each treemap element, corresponding to a tree node, is a dynamic object that enables the user to access the full information of the instance represented by the node, including of all the values of the frailty variables.

Figure 5 shows the treemap view of the tree structure presented in Fig. 4. The intensity of the colour used to represent each node in the treemap depends on the degree of similarity in relation to the parent node. Stronger intensity levels indicate a higher degree of similarity between the parent and child nodes.

We have implemented a procedure to represent the treemap view of the similarity results in the Android mobile application. This procedure is known as the affinity result task and is accessed from the main application dashboard. Figure 6 shows a screenshot of the results on a tablet device. The physician can access the detailed view of a specific patient instance by touching the screen item associated with the instance node (in this case, the selected node refers to a specific patient instance).

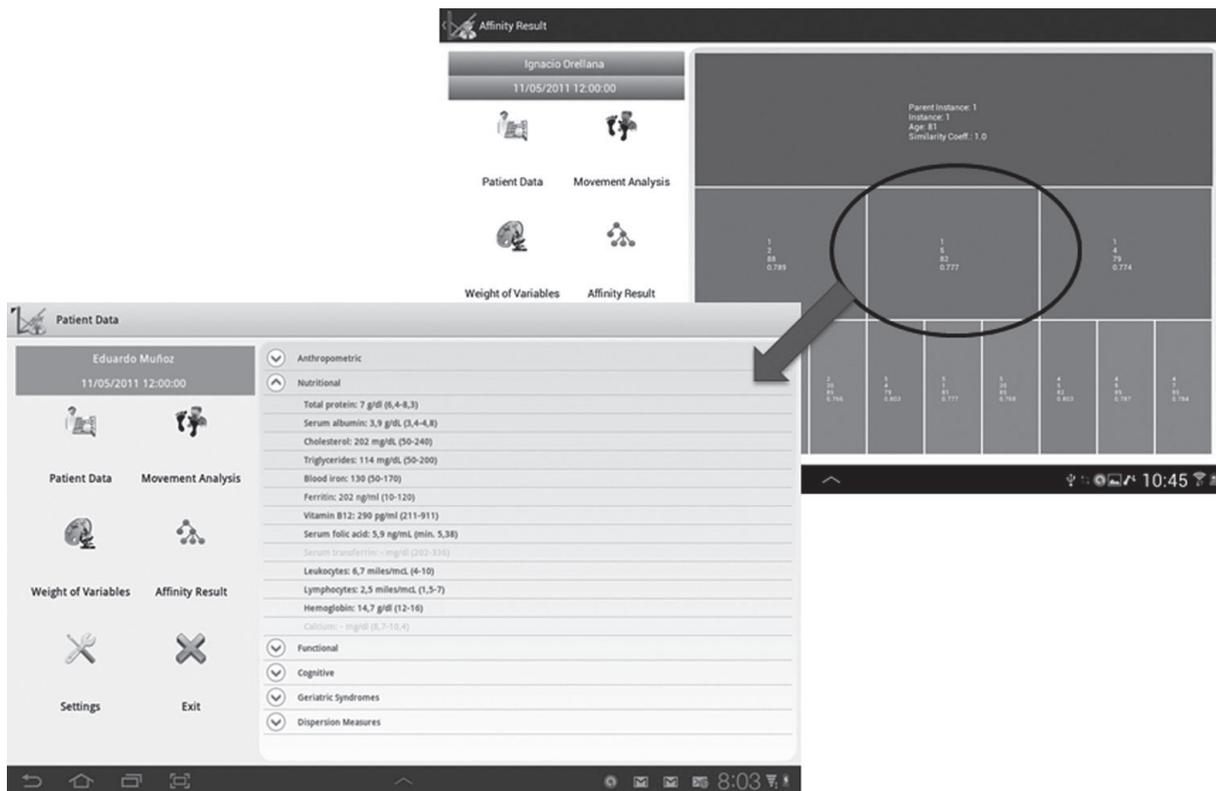


Fig. 6. Affinity result option in the mobile application. Screenshots of the treemap view for a specific patient and a detailed view of frailty variable values for a selected treemap item are shown.

5.5.1. Integration in a frailty mobile application

We have developed the proposed service as a web service hosted on a server. The service can therefore be included as a component of other systems, improving its scalability. Using a service-oriented approach, complex tasks are delegated to a server, and the results are obtained through service. In our case, examples of the processes that may be implemented include the frailty risk variables from the patient record, adjustment of the variable weights and calculation of the similarity results.

The final frailty assessment, represented in the treemap view, is integrated into the mobile application to provide the physician with a visual mechanism to aid in the final diagnosis for the selected patient. Other additional features are also provided by the mobile application, as mentioned broadly in previous sections. A summary of these features is as follows:

- Gathering and calculating dispersion measures using accelerometer-embedded mechanisms.
- Collecting frailty variables from the patient record.
- Displaying the state of the elder based on the measures provided by these variables.
- Editing the weights of the frailty variables and modifying the tree depth.
- Selecting the minimum similarity coefficient and main colours of the treemap structure.
- Displaying the final treemap of results for diagnosis support.

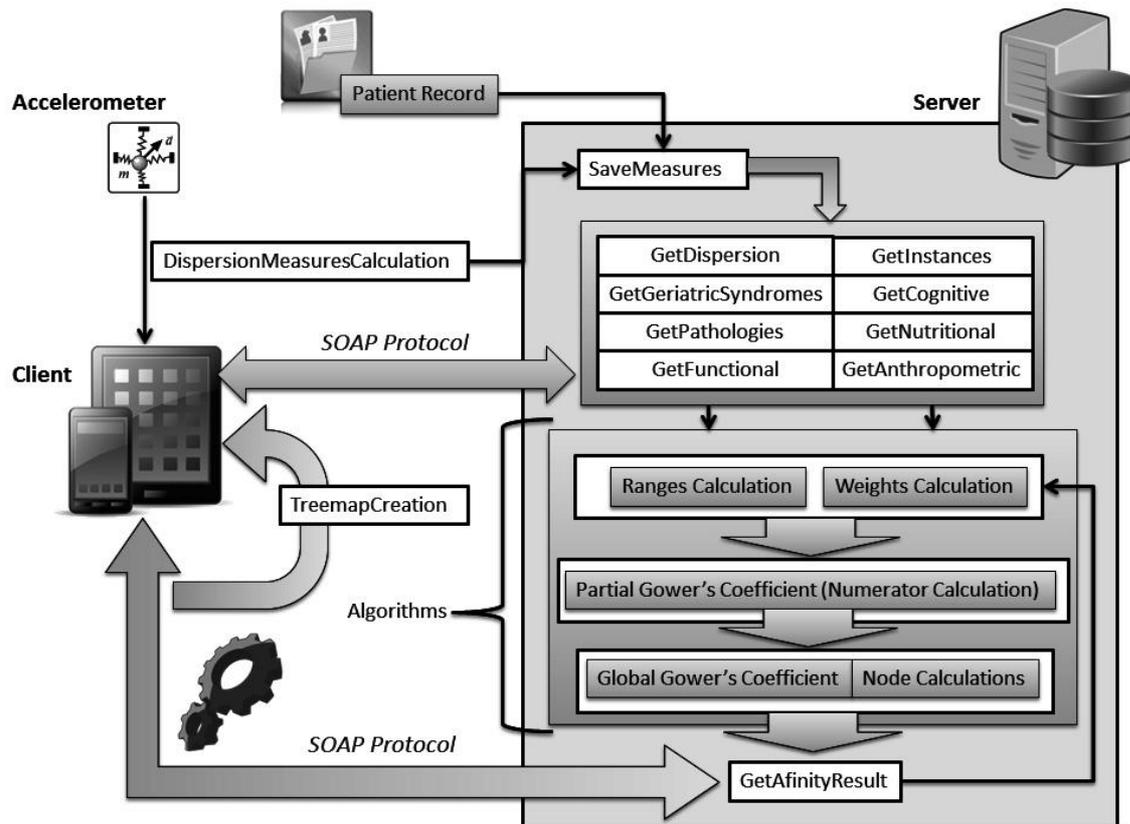


Fig. 7. Infrastructure overview. Elements, algorithms and web services.

The independence of the various processes described in this article, such as the calculation of the Gower coefficients, is important for their integration into other applications as software components.

5.6. General infrastructure. Elements and features

The web services infrastructure has been deployed in a real scenario considering the schema presented in Fig. 7. Three hardware elements have been identified:

- Accelerometer. We have used the mobile phone's accelerometer to collect data from gait activity. However, the mobile device could communicate with an accelerometer device via Radiofrequency or Bluetooth, sending the movement data to the mobile device.
- Mobile Device. The developed frailty mobile application is deployed in Tablet devices and/or mobile phones. These devices communicate with the server through web services, and they carry out the functions mentioned in Section 5.5.1.
- Server. The database and the web services are hosted in a server. In this case, MySQL has been used as relational database management system and Internet Information Services (IIS) has been used as web server. Thus, the web services have been implemented by using the Microsoft .NET

Table 3
Descriptive excerpt of selected patient instances

Anthropometric								
Sex	Size (cm)	Weight (Kg)	BMI (Kg/m ²)	Body mass (Kcal)	Fat mass (%)	Lean mass (Kg)	Total water (Kg)	
Female	159,6 ± 6,02	65,45 ± 11,36	28,14 ± 4,06	1158,7 ± 122,9	36,37 ± 8,57	40,93 ± 6,62	31,11 ± 5,22	
Male	150,9 ± 5,03	67,78 ± 8,46	26,19 ± 2,7	1246,7 ± 128,05	27,61 ± 4,34	48,29 ± 4,13	31,11 ± 5,22	
Nutritional (relevant biochemical values)								
Sex	Hemoglobin (g/dl)	Ferritin (ng/ml)	Total protein (g/dl)	Serum albumin (g/dl)	Cholesterol (mg/dl)	Serum folic acid (ng/ml)	Vitamin B12 (pg/ml)	Lymphocytes (thousands/mcL)
Female	12,92 ± 1,13	81,91 ± 70,69	6,95 ± 0,6	4,18 ± 0,29	173,55 ± 31,46	7,42 ± 4,54	354,73 ± 70,08	1,71 ± 0,67
Male	14,57 ± 0,88	165,22 ± 143,92	6,88 ± 0,44	3,95 ± 0,2	177,89 ± 29,59	6,32 ± 3,22	397,56 ± 109,19	1,9 ± 0,56
Geriatric Syndromes (11)		Pathologies			Functional			
Sex	Average	Sex	Average	Sex	Barthel Index	Tinetti gait score	Tinetti balance score	Lawton and Brody score
Female	4,57	Female	3,9	Female	95 ± 4,90	11,45 ± 0,78	13,82 ± 1,53	6 ± 1
Male	2,29	Male	3	Male	99,9 ± 0,3	12 ± 0	14,7 ± 1,1	4 ± 0

Table 4
Node List calculated for instance 1 (second level)

Parent Inst. Id.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Instance Id.	12	16	2	15	4	9	7	5	20	18	19	17	13	10	11	6
Name	P12	P16	P2	P15	P4	P9	P7	P5	P20	P18	P19	P17	P13	P10	P11	P6
Age	85	87	88	81	79	89	85	82	85	87	80	73	82	82	92	89
Similarity Coeff.	0,734	0,72	0,716	0,715	0,709	0,705	0,702	0,702	0,702	0,683	0,679	0,671	0,658	0,64	0,63	0,623

framework. SOAP² was the communication protocol used to establish the exchange of information between consumer and producer of web services.

All the web services are identified by the prefix “Get”. The mobile device requests the value of frailty measures by means of the corresponding “get” call (e.g., GetFunctional or GetCognitive). When the mobile device asks for the frailty results the “GetAffinityResult” web method is called, and the following algorithms are run: “RangesCalculation”, “WeightsCalculation”, “PartialGowerCoefficient”, “GlobalGowerCoefficient” and “NodesCalculation”. These algorithms use the results from the “Get” methods internally. The response of the “GetAffinityResult” web method contains the final nodelist to be used by the mobile device to generate the treemap.

6. Experimental results

In this section, we present the results provided by the developed service for an initial cluster of sixty instances from 20 elderly patients with equal numbers of women and men. The average age of the men is $81,8 \pm 4,74$ while the women have an average age of $85,45 \pm 3,22$. All of the patients have adequate social abilities and are independent in their everyday activities, and none of them exhibit cognitive impairment.

²http://www.w3.org/standards/techs/soap#w3c_all.

Table 5
Node List calculated for instances 12, 16, 2 (third level)

Parent Inst. Id.	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
Instance Id.	15	16	1	13	7	4	18	9	8	5	11	2	17	19	20
Name	P15	P16	P1	P13	P7	P4	P18	P9	P8	P5	P11	P2	P17	P19	P20
Age	81	87	81	82	85	79	87	89	85	82	92	88	73	80	85
Similarity Coeff.	0,751	0,75	0,734	0,713	0,707	0,689	0,682	0,679	0,67	0,656	0,655	0,649	0,649	0,648	0,634
Parent Inst. Id.	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
Instance Id.	18	12	1	2	15	19	13	9	8	17	11	4	5	7	6
Name	P18	P12	P1	P2	P15	P19	P13	P9	P8	P17	P11	P4	P5	P7	P6
Age	87	85	81	88	81	80	82	89	85	73	92	79	82	85	89
Similarity Coeff.	0,773	0,75	0,72	0,719	0,718	0,706	0,705	0,7	0,699	0,692	0,683	0,679	0,671	0,665	0,662
Parent Inst. Id.	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Instance Id.	16	1	20	17	18	4	5	9	15	6	12	19	10	8	
Name	P16	P1	P20	P17	P18	P4	P5	P9	P15	P6	P12	P19	P10	P8	
Age	87	81	85	73	87	79	82	89	81	89	85	80	82	85	
Similarity Coeff.	0,719	0,716	0,699	0,697	0,681	0,671	0,653	0,652	0,651	0,65	0,649	0,64	0,629	0,62	

Table 6
Node List calculated for instance 1, from anthropometric and nutritional viewpoint ($S \geq 0.62$)

Parent Inst. Id.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Instance Id.	18	5	10	20	15	2	19	7	16	4	12	17	9	6	8
Name	P18	P5	P10	P20	P15	P2	P19	P7	P16	P4	P12	P17	P9	P6	P8
Age	87	82	82	85	81	88	80	85	87	79	85	73	89	89	85
Similarity Coeff.	0,782	0,756	0,736	0,735	0,729	0,725	0,715	0,714	0,713	0,71	0,709	0,684	0,663	0,632	0,628

First, the values of the frailty risk variables (or at least most of them) must be obtained for each patient in the group. For each elder, we collected measures of the variables at three different times and created the corresponding instances. However, we will consider the values of the first instance for each patient to understand the results provided by the service more easily. Table 3 shows a descriptive overview of the first dataset of instances, with some of the values of the frailty variables.

The Android mobile application was used during gait activity for collecting accelerometer data and calculating the related dispersion measures. Note that when a new patient is studied, a new instance is created and stored in the system, improving the accuracy of the subsequent set of similarity results.

For a given geriatrician, a frailty assessment can be performed based on four different criteria, anthropometric, functional, nutritional and gait-related, where the importance assigned to each variable in each domain can vary. In this case, we have assigned the maximum importance to the entire group of variables (100%). Therefore, if we focus on the frailty assessment procedure for instance 1, a node list sorted by the similarity coefficient is created, but only the first three nodes form the second level of the treemap (the first corresponding to the root instance, i.e., instance 1). Table 4 shows the results of Gower’s calculation between instance 1 and the remaining instances. The columns in bold represent the nodes that contribute to the tree. In this case, instances with similarity values lower than $S = 0,62$ have been ignored because the expert determined that lower similarity values would not be as useful in making decisions.

Instance 1 has a similarity degree of 73.4% with instance 12 corresponding to Patient 12, 72% with instance 16 and 71.6% with instance 2. For the three level treemap, the Gower coefficients between these instances and the remainder of the instances are calculated (Table 5 shows the results). The children of the last node (corresponding to instance 2) exhibit lower similarity coefficients of 71.9%, 71.6% and 69.9%.

These results aid the physician in determining the frailty condition of the specified patient at a given moment, in relation to other patients in the same elderly population.

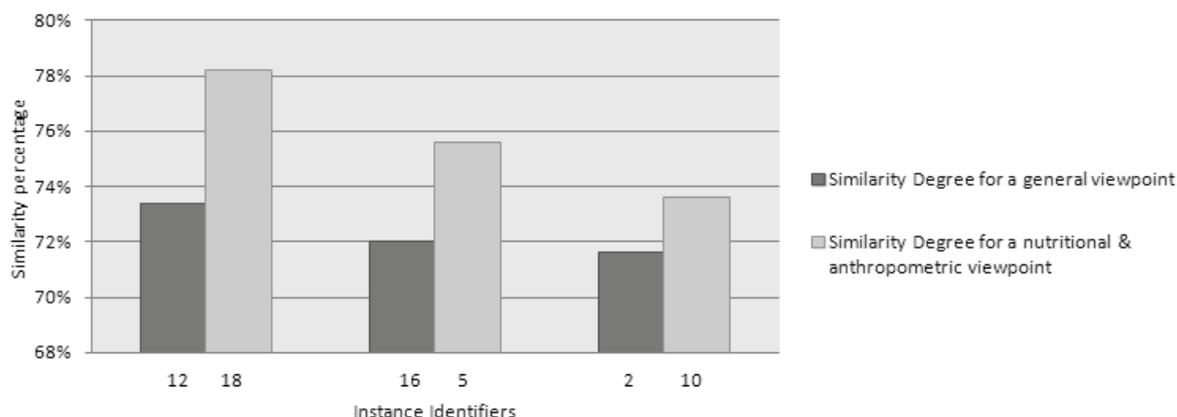


Fig. 8. Comparison of similarity degrees for instance 1 (Patient 1).

However, the behaviour of the algorithm and the final results will be different if the weights values of the frailty variables are modified using the mobile application user interface. In this case, we consider only the anthropometric and nutritional variables (with 100% importance) and ignore the remainder of the domains (setting their weights to 0%). Table 6 shows the node list for instance 1 with these new requirements.

Figure 8 shows a comparison chart between the nodes of the second level of the treemap from a global viewpoint (considering all domains) and a more narrow nutritional and anthropometric viewpoint for instance 1 for Patient 1. The similarity results and instances are different because of the significance of the analysis. The similarity degrees are higher when the study is focused on specific domains rather than a global analysis. However, the elderly patients are studied from a particular perspective which is more appropriate under some circumstances (e.g., for adults without nutritional disorders and cognitive impairments, analysing only the functional domain and gait). Based on these results, the doctor can group the individuals according to their similarity degrees for optimal care and treatment. From the viewpoint of performance, the maximum time to generate both treemaps on the mobile screen were 2.55 and 2.19 seconds respectively (working on sixty instances from the twenty patients); we observed less processing time when a shorter group of variables was studied.

7. Conclusions and future work

In this work, we have presented a novel service infrastructure for supporting clinical decisions regarding frailty in the elderly. The aim of this service is to calculate an index or similarity coefficient that accounts for all relevant frailty variables in the record of the elderly patient. Owing to the mixed nature of variables under study (quantitative, qualitative and binary), Gower's algorithm is the most appropriate method for obtaining the similarity values for a group of patient instances. The obtained results are transformed into programming objects known as nodes which are represented using a treemap structure based on the similarity values calculated for each node.

This initial approach has been deployed in a ubiquitous environment. We have developed an internal web services infrastructure that communicates with a mobile frailty application in which the results are shown in a treemap view. Additionally, this approach can be deployed on other devices (not only mobile

ones), especially for better results visualization (e.g., widescreens); however, the mobile device is also used due to the heterogeneity and mobility of the clinical environments, getting evaluations at any time (when the physician needs them). The mobile application makes it possible to determine the values of the frailty risk factors for a specific elderly patient and to configure different parameters such as the variable weights or tree depth to adapt the analysis to a specific domain (nutritional, anthropometric, functional, cognitive, etc.) with particular settings. This approach allows the doctor to focus on clinical areas in which the patient requires more care. We also used existing mobile accelerometer devices to analyse the functional characteristics of the patients, using the gait activity of the patient as a primary frailty indicator as proposed in other related studies. We propose a new group of variables known as dispersion measures, that are calculated by the mobile application from the gait activity. These dispersion measures are incorporated into the similarity calculation.

A broad group of patient instances must be stored in the system prior to the deployment of the proposed service. As the number of patient instances and stored variables increases, the system yields increasingly accurate results, aiding physicians in determining the frailty conditions of elders. We have demonstrated that the accuracy of the results depends on the number of variables with an associated value. Therefore, if the patient instance under study includes empty variables, then these variables are not taken into account in the similarity calculation, and the results are therefore incomplete. This limitation owes to the unavailability of clinical information in certain cases, which affects some of the information sought. The creation of protocols and clinical mechanisms for collecting the values of the related frailty factors may help to resolve this issue. We could not make a comparison with current methods because these are not standardised, depend on the physician experience and they would not provide reliable results at this moment. Additionally, the number of geriatricians involved in the experiment was limited to afford feasible conclusions in this aspect.

In future work, we consider to perform an study of the usability of the mobile applications as well as the acceptance level of the whole system. Additionally, we hope to integrate the service developed in this paper into a more complex and distributed system, accounting for various clusters and populations of elderly people. This goal requires a large set of patient instances, including a broad range of patient ages to determine the appropriate degree of membership to a specific cluster from an initial cluster analysis prior to the application of the Gower similarity service.

Acknowledgments

This work was partially funded by the TIN2010-20510-C04-04 project from Ministerio de Ciencia e Innovación (Spain). In addition, we sincerely appreciate the clinical support of Fco. Javier Navarro from the geriatric services department, Residencia Asistida de Ancianos, Ciudad Real, Spain.

References

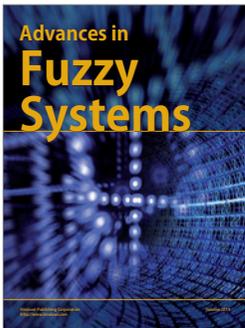
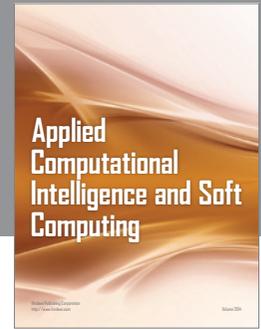
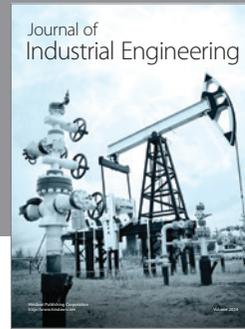
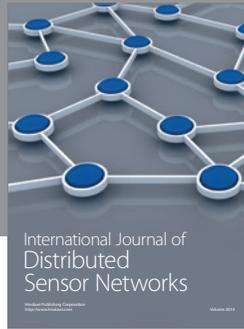
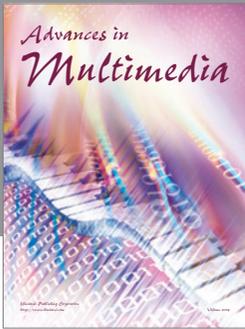
- [1] F.I. Mahoney and D. Barthel, Functional evaluation: the Barthel Index, *Maryland State Medical Journal* **14** (1965), 56–61.
- [2] S. Espinoza and L. Fried, Risk Factors for Frailty in the Older Adult, *Clinical Geriatrics* **15**(6) (2007), 37–44.
- [3] B. Everitt, Cluster Analysis, 2th edition, McGraw-Hill bokk Company, N. York, 1980.
- [4] J. Fontecha, F.J. Navarro, R. Hervás and J. Bravo, Elderly frailty detection by using accelerometer-enabled smartphones and clinical information records, *Personal and Ubiquitous Computing Journal*, Springer DOI: 10.1007/s00779-012-0559-5, 2012.

- [5] L.P. Fried, J. Waltson and A.B. Newman, Frailty in older adults: evidence of a phenotype, *J Gerontology* **56**(A)(3) (2001), 146–156.
- [6] E. Gelnarová and L. Safarík, Comparison of three statistical classifiers on a prostate cancer data, *Neural Network World* **15**(4) (2005), 311–318.
- [7] M.R. Gillick, Long-term care options for the frail elderly, *J Am Geriatr Soc* **37** (1989), 1198–1203.
- [8] R.J. Gobbens, K.G. Luijckx, M.T. Wijnen-Sponselee and J.M. Schols, Towards an Integral Conceptual Model of Frailty, *The Journal of Nutrition Health and Aging* **14** (2010), 175–181.
- [9] L.S.A. Gonçalves, R. Rodrigues, A.T. Amaral Júnior, M. Karasawa and C.P. Sudré, Comparison of multivariate statistical algorithms to cluster tomato heirloom accessions, *Genetics and Molecular Research* **7**(4) (2008), 1289–1297.
- [10] A.D. Gordon, Cluster Classification, Wiley New York, 1990.
- [11] J. Gower, A general coefficient of similarity and some of its properties, *Biometrics* **27** (1971), 857–872.
- [12] D. Hamerman, Toward an understanding of frailty, *Ann Intern Med* **130** (1999), 945–950.
- [13] P. Jaccard, Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines, *Bulletin de la Société Vaudoise des Sciences Naturelles* **37** (1901), 241–272.
- [14] D.M. Jones, X. Song and K. Rockwood, Operationalizing a frailty index from a standardized comprehensive geriatric assessment, *J Am Geriatr Soc* **52** (2004), 1929–1933.
- [15] T.M. Lim and H.W. Khoo, Sampling properties of Gower's general coefficient of similarity, *Ecology* **66**(5) (1985), 1682–1685.
- [16] G. Londoño, L. Lavalett, P. Galindo and L. Afanador, Use of multivariate methods for grouping strains of *Colletotrichum* spp. based on cultural and morphological characters, *Rev Fac Nal Agr Medellín* **60**(1) (2007), 3671–3690.
- [17] H. Pinnock, R. Slack, C. Pagliari, D. Price and A. Sheikh, Understanding the potential role of mobile phone-based monitoring on asthma self-management: Qualitative study, *Clin Exp Allergy* **37**(5) (2007), 794–802.
- [18] S.D. Searle, A. Mitnitski, E.A. Gahbauer, T. Gill and K. Rockwood, A standard procedure for creating a frailty index, *BMC Geriatrics* **8** (2008), 8–24.
- [19] B. Shneiderman, Treemaps for space-constrained visualization of hierarchies at <http://www.cs.umd.edu/hcil/treemap-history/index.shtml>, last update June 25th, 2009 by Catherine Plaisant, last access November 5th, 2012.
- [20] S.S. Socransky, A.C. Tanner, J.M. Goodson, A.D. Haffajee et al., An approach to the definition of periodontal disease syndromes by cluster analysis, *Journal of Clinical Periodontology* **9**(6) (1982), 460–471.
- [21] A.J. Soto, I. Ponzoni and G.E. Vázquez, Numerical Analysis of Different Similarity Criteria on Clustering Algorithms, XV Congress on Numerical Methods and their Applications – ENIEF 2006. November 7–10, 2006.
- [22] P.-N. Tan, M. Steinbach and V. Kumar, Introduction to Data Mining – Ch.8: Cluster Analysis: Basic Concepts and Algorithms, Addison-Wesley, 2006.
- [23] M.E. Tinetti, Performance-oriented assessment of mobility problems in elderly patients, *Journal of the American Geriatrics Society* **34**(2) (1986), 119–126.
- [24] V. Villarreal, J. Laguna, S. López et al., Proposal for mobile diabetes self-control: Towards a patient monitoring framework, *Proceedings of the 10th International Work-Conference on Artificial Neural Networks-IWANN'09* **5518** (2009), 870–877.
- [25] K.W. Woodhouse, H. Wynnie, S. Baillie et al., Who are the frail elderly? *Q J Med* **68** (1988), 505–506.

Jesús Fontecha Diezma graduated from the University of Castilla-La Mancha in computer engineering in September 2008. In 2010, he received his M.S. degree in advanced computer technologies from Castilla-La Mancha University. He is currently enrolled as a PhD student at the same University, performing research in the Modelling Ambient Intelligence Research Group (MAmI), with a primary focus on frailty, mobile computing, AAL and gait analysis. In the past years, he has participated in several healthcare projects in collaboration with other universities and companies. His research interests also include ubiquitous computing, mobile sensors, context awareness, data mining and visualisation on ambient displays.

Ramón Hervás Lucas is an associate professor of computer science in the Department of Languages and Information Systems at the Castilla-La Mancha University, Spain and a member of the Modelling Ambient Intelligence Research Group (MAmI). He received his M.Sc. and Ph.D. degrees in computer engineering from the UCLM University. He has been invited as a lecturer at various universities and collaborates on several research projects. Dr. Hervás research interests cover a variety of IT issues including context-awareness, ubiquitous computing, semantic web, intelligent user interfaces and information visualisation.

José Bravo Rodríguez is an associate professor of computer science in the Department of Languages and Information Systems at Castilla-La Mancha University, Spain and the director of the Modelling Ambient Intelligence Research Group (MAmI). Since 2003, Dr. Bravo has organised the International Conference on Ubiquitous Computing and Ambient Intelligence (UCAmI) and the International Conference on Ambient Assisted Living (IWAAL). He is involved in research in a variety of areas, including ubiquitous computing, ambient intelligence, ambient assisted living, context-awareness and mobile computing.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

