Project Title: **Sensing and predictive treatment of frailty and associated co-morbidities using advanced personalized models and advanced interventions**

Contract No: 690140

Instrument: Collaborative Project

Call identifier: H2020-PHC-2014-2015

Topic: PHC-21-2015: Advancing active and healthy ageing with ICT: Early risk detection and intervention

Start of project: 1 January 2016

Duration: 36 months

Deliverable No: D4.14

Signal processing algorithms for extraction of frailty related indicators (Preliminary version)

Due date of deliverable: M12 (31st December 2016)

Actual submission date: 31st December 2016

Version: 1.0

Lead Authors: Konstantinos Deltouzos, Aimilia Papagiannaki Evangelia Pippa, Dimitrios Vlachakis, Vasileios Megalooikonomou (UoP), Andreas Vasilakis, Konstantinos Moustakas (CERTH)

Lead partners: UoP, CERTH

,

Horizon 2020
European Union funding
for Research & Innovation

Change History

Ver.	Date	Status	Author (Beneficiary)	Description
0.1	24/11/2016	Draft	Dimitrios Vlachakis (UoP)	First draft from WP leader
0.2	6/12/2016	Draft	Konstantinos Deltouzos (UoP), Evaggelia Pippa (UoP), Yiannis Ellul (UoP), Andreas Vasilakis (CERTH), Ilias Kalamaras (CERTH)	First round of contribution collected
0.3	6/12/2015	Draft	Konstantinos Deltouzos (UoP)	Feedback by deliverable responsible: 1st draft sent out to all partners.
0.4	15/12/2016	Draft	Athanassios Kordas (UoP), Evaggelia Pippa (UoP), Aimilia Papagiannaki (UoP), Andreas Vasilakis (CERTH), Konstantinos Moustakas (CERTH)	Second round of contribution collected
0.5	20/12/2016	Draft	Konstantinos Deltouzos (UoP)	Feedback by deliverable responsible: 2nd draft sent out to all partners.
0.6	22/12/2016	Draft	Aimilia Papagiannaki (UoP), Vasileios Megalooikonomou (UoP)	Incorporation of additional input
0.7	23/12/2016	Draft	Konstantinos Deltouzos (UoP), Dimitrios Vlachakis (UoP)	Several revisions taking into account internal review's comments
0.8	27/12/2016	Draft	Aimilia Papagiannaki (UoP), Vasileios Megalooikonomou (UoP)	Additional contributions
1.0	28/12/2016	Final	Dimitrios Vlachakis (UoP)	Deliverable finalized.

EXECUTIVE SUMMARY

The aim of work package **WP4** is to develop methods for the offline and online management, fusion and analysis of multimodal and advanced technology data from social, behavioral, cognitive and physical activities of frailty older people and apply them to manage and analyze new data. Results from the analysis of existing and new data will be also used to create user-profiling virtual models of elderly patients. Towards this direction, the task T4.6 aims to examine methods of fusing information to extract frailty related indicators. These methods need to manage uncertainty in the system generated by incompleteness and noise of wearable sensor data.

In this deliverable, our primary efforts were focused on identifying approaches for discovering a set of relevant and informative indicators for frailty. During this process, the state of the art was first analyzed and the clinical experts of our consortium gave their valuable input. Then, the multidimensional data analysis problem was formulated using tensors as a tensor decomposition problem and several techniques were outlined. Moreover, preliminary work was performed on data mining techniques towards discovering associations between frailty, and physiological or behavioral patterns. These techniques aim to discover multi-level association rules, in a distributed environment, from multiple heterogeneous data sources. Finally fueled by previous work on data fusion, three schemes were explored: (i) Early Integration scheme, (ii) Late Integration scheme with local (sensor dependent) training models, (iii) Late Integration scheme with global (sensor independent) training model. This first version of the deliverable whose final version is due on M24 sets the ground for the data analysis techniques that will be used to discover new frailty metrics.

DOCUMENT INFORMATION

Contract Number:	H2020-PHC–690140	**Acronym:**	FRAILSAFE
Full title	Sensing and predictive treatment of frailty and associated co-morbidities using advanced personalized models and advanced interventions		
Project URL	http://frailsafe-project.eu/		
EU Project officer	Mr. Jan Komarek		

Deliverable number:	4.14	**Title:**	Signal processing algorithms for extraction of frailty related indicators
Work package number:	4	**Title:**	Data Management and Analysis

Date of delivery	**Contractual**	31/12/2016 (M12)	**Actual**	31/12/2016
Status	Draft ☒		Final ☐	
Nature	Report ☒	Demonstrator ☐	Other ☐	
Dissemination Level	Public ☐	Consortium ☒		
Abstract (for dissemination)	In this deliverable, our primary efforts were focused on identifying approaches for discovering a set of relevant and informative indicators for frailty. During this process, the state of the art was first analyzed and the clinical experts' input was taken into account. Then, the multidimensional data analysis problem was formulated as a tensor decomposition problem and several techniques were outlined. Moreover, preliminary work was performed on data mining techniques towards discovering associations between frailty, and physiological or behavioral patterns. Finally, fueled by previous work on data fusion, three schemes were explored. This first version of the deliverable whose final version is due on M24 sets the ground for the data analysis techniques that will be used to discover new frailty metrics.			
Keywords	FrailSafe, signal processing, frailty indicators, tensors, data fusion			

Contributing authors (beneficiaries)	Konstantinos Deltouzos, Evangelia Pippa, Dimitrios Vlachakis, Aimilia Papagiannaki, Vasileios Megalooikonomou (UoP), Andreas Vasilakis, Konstantinos Moustakas (CERTH)			
Responsible author(s)	Konstantinos Deltouzos, Vasileios Megalooikonomou		**Email**	deltouzos@upatras.gr
	Beneficiary	UoP	**Phone**	+30 2610 996994

TABLE OF CONTENTS

TABLE OF FIGURES 5
LIST OF TABLES 6
LIST OF ABBREVIATIONS AND ACRONYMS 6
1 INTRODUCTION 7
2 FRAILTY INDICATORS 8
2.1 Frailty phenotype and Frailty Index 8
2.2 Electronic Frailty Index (eFI) 10
2.3 Definition of Frailty Indices in FrailSafe 12
3 SIGNAL PROCESSING AND DATA MINING TECHNIQUES FOR EXTRACTING FRAILTY RELATED INDICATORS 13
3.1 Analysis using tensors 13
3.1.1 Tensor decomposition 14
3.1.2 Applying tensors to FrailSafe data 18
3.2 Mining of multi-level association rules 19
3.2.1 Architectural Model 20
3.2.2 Current and future work 22
4 DATA FUSION 24
REFERENCES 27

TABLE OF FIGURES

Figure 1: Map of the 36 deficits used in the eFI 11
Figure 2: Third order tensor. 13
Figure 3: Slices and fibers of a third order tensor 14
Figure 4: Tucker decomposition of a third order tensor 14
Figure 5: PARAFAC decomposition of a third order tensor 14
Figure 6: Classification diagram based on TUCKER decomposition 17
Figure 7: Constrained CMTF model 19
Figure 8: Multi-tier model 21
Figure 9: System control flow diagram 22
Figure 10: Feature-level fusion scheme. 24
Figure 11: Decision-level with local (sensor dependent) training models fusion scheme. 25

Figure 12: Decision-level with global (sensor independent) training model fusion scheme. 26

LIST OF TABLES

Table 1: Main characteristics of the frailty phenotype and the Frailty Index 8
Table 2: List of the 36 deficits used in the eFI 11
Table 3: Frailty Indices definition 12

LIST OF ABBREVIATIONS AND ACRONYMS

(*in alphabetic order*)

ALS	Alternating Least Squares
CANDECOMP	CANonical DECOMPosition
ClFI	Clinical Frailty index
CoFI	Combined Frailty index
ECG	Electrocardiography
eFI	Electronic Frailty Index
HER	Electronic Health Records
EI	Early Integration
HOSVD	High-Order Singular Value Decomposition
LI	Late Integration
LSA	Latent Semantic Analysis
PARAFAC	PARAllel FACtor analysis
PCA	Principal Components Analysis
SSH	Secure Shell
SVD	Singular Value Decomposition
TFI	Technical Frailty Index
XML	eXtensible Markup Language

1 INTRODUCTION

One of the key objectives of the FrailSafe project is the better understanding of frailty and the development of new quantitative and qualitative measures to define it. Towards this direction, the current state of the art of frailty definition was analyzed in order to discover the strengths and limitations of each method. In Chapter 2 a summary of this work is presented, and a special attention is given to the electronic Frailty Index which can be generated automatically by health record data. During this process, the clinical experts of our consortium gave their valuable input.

In Chapter 3 we present some preliminary work on signal processing and data mining techniques. This work is focused so far in two directions. First is the modelling of the multimodal data that are being collected in FrailSafe, using tensors. By doing so, we can use their strong mathematical background and achieve several advantages such as data compression, identification of clusters and exploitation of patterns. The second direction aims to discover a novel way of mining multi-level association rules, in a distributed environment, from multiple heterogeneous data sources. The general architecture model and the future work on this are presented in Section 3.2 .

Finally in Chapter 4 our preliminary work in designing data fusion schemes is presented. Several different approaches for fusing data from different sensor units/dimensions are being explored, resulting in three proposed schemes: (i) Early Integration scheme, (ii) Late Integration scheme with local (sensor dependent) training models, (iii) Late Integration scheme with global (sensor independent) training model. These schemes will be tested and validated in the following period with the FrailSafe sensor data.

2 FRAILTY INDICATORS

2.1 Frailty phenotype and Frailty Index

The frailty syndrome has been widely discussed among the scientific community and its foundations are generally well established in the literature. However its practical interpretation, particularly in the ordinary clinical practice, remains questionable [1]. The combination of frailty measures in clinical practice is essential for the mediations and interventions design against age-related conditions (such as disability) in older people [2]. Several methods have been developed lately in an effort to address this geriatric multidimensional syndrome.

The authors in [3], focused initially at some basic clinical manifestations of frailty, which were then projected into the frailty phenotype as it was described in the Cardiovascular Health Study [4]. Along these lines, Rockwood et al. [5] utilized the Canadian Study of Health and Aging to create and approve their proposed Frailty Index. Additional methods to quantify frailty have been proposed over the last years, mainly expanding on these two models [6] [7] [8]. Indeed, the frailty phenotype and the Frailty Index are considered by the scientific community as the cornerstone of frailty definition. These two methods follow a different approach, and thus should be considered complementary [9]. Their main characteristics are summarized in Table 1.

Frailty phenotype	Frailty Index
Signs, symptoms	Diseases, activities of daily living, results of a clinical evaluation
Possible before a clinical assessment	Doable only after a comprehensive clinical assessment
Categorical variable	Continuous variable
Pre-defined set of criteria	Unspecified set of criteria
Frailty as a pre-disability syndrome	Frailty as an accumulation of deficits
Meaningful results potentially restricted to non-disabled older persons	Meaningful results in every individual, independently of functional status or age

Table 1: Main characteristics of the frailty phenotype and the Frailty Index

The frailty phenotype uses five distinct criteria that assess the appearance of signs or manifestations related to frailty (involuntary weight loss, slow walking speed, poor handgrip quality, reported exhaustion and mobility issues) [4]. The quantity of criteria being met by the subject leads to a 6-level ordinal variable extending from 0 to 5. This is then sorted into a 3-level variable portraying a fit older person (none of the criteria), a pre-frail person (meets one or two criteria) and a frail person (meets at least three criteria). The frailty phenotype can be performed at the first meeting with the subject and does not require an in depth clinical assessment. In this way, it serves as a general categorization of the population into three distinct profiles. Overall, the frailty phenotype does not give any specific guidelines about preventive or helpful mediations or interventions. The problem is that it is composed of extremely broad signs or side effects, which are only able to raise an alert about a potential health issue. This is not enough though to design a quick preventive or restorative intervention due to the fact that there is no information about the underlying cause of frailty. For instance, it is clinically impossible to treat sudden weight loss or slow gait speed without knowing the basic causal conditions. This is only possible by a thorough geriatric clinical assessment, in which the overall health status of the older person is being assessed through a multidimensional, interdisciplinary analytic process leading to particular clinical interventions.

The Frailty Index is made by a long list of clinical conditions, disorders, and diseases. There are more than 70 parameters as an initial screening tool that must be addressed. The actual goal of this list is to address and bring to the surface more critical deficiencies that have accumulated over years. Although the Frailty Index has many times been revised and updated, the end goal remains to be able to clearly reflect on dichotomous conditions (e.g. robustness versus frailty). It is clear that the Frailty Index is impractical and inapplicable as a first contact tool for frailty diagnosis, since an extensive geriatric assessment of the older person must be conducted at the same time. Once the full assessment is complete, the Frailty Index can be used as a tool for monitoring the continuous follow up of the older person. Actually, the Frailty Index is more sensitive to changes than the overall frailty phenotype. Consequently, the Frailty Index might be of more use to help the clinician determine the effectiveness of any intervention that was designed and to depict the health status progress of the individual over time. In any case, the clinical intervention dependably goes through the Frailty Index's categorization into classes of frailty, separating normal ageing from anomaly. The categorization into risk groups of the frailty phenotype makes it more powerful as a tool that will link a typical clinical condition to frailty. In a clinical world that is constantly dominated by new advancements and developments, it can be of great value to formulate a complete geriatric assessment tool that can be generated by an electronic health record and serve as a reference for following assessments.

To sum up, it can't be overlooked that there are two noteworthy theoretical contrasts at the heart of the two frailty assessment tools:

(1) Relationship amongst frailty and age-related grouped conditions. As said, the frailty phenotype depends on the assessment of signs and clinical manifestations. This implies, as indicated by Fried et al. [4] [10] that frailty may hypothetically exist even without medically characterized conditions. Under such viewpoint, the frailty phenotype for sure portrays a novel age-related field of research for medical sciences [11]. Then again, the Frailty Index revolves to a great extent around medical grouped conditions. It depicts a likeliness profile that is close to the one assessed by the clinician, which is possibly capable to characterize the phenotype frailty and to link it to its early signs as a preventive tool.

(2) Relationship of frailty with disability. In their review assessing the phenotype, Fried et al. [4] suggests that frailty causes disability that may not be linked to (sub)clinical disorders. They clarify that 'the syndrome of frailty may be a physiologic precursor and etiologic factor in disability'. This implies a verifiable identification of frailty as a key element for the design and conduction of interventions against episodes that may result in disability. Along these lines, the frailty phenotype finds its optimal application in non-disabled more independent subjects. Then again, the Frailty Index incorporates measures of everyday incapacity (e.g. issues with getting dressed, issues with washing and reduced versatility) in its calculation [5]. At the end of the day, the Frailty Index does not make a clear distinction between frailty and disability. It is more focused at impartially evaluating the measure of accumulated deficits of each individual, whichever they are.

These conceptual differences between the two instruments obviously and consequently differentiate the target populations to which they might be applied. As mentioned individual, while we may meaningfully estimate the Frailty Index in every the frailty phenotype may lose some of its clinical relevance when assessed in older persons already experiencing disability.

To summarize, the frailty phenotype categorically defines the presence/absence of a condition of risk for subsequent events (most specifically, disability). By differentiating a normal (i.e. robustness) versus an abnormal (i.e. frailty) status, the frailty phenotype may facilitate the implementation of the frailty concept into clinical practice. It provides the clinical-friendly dichotomous variable on which deciding the possible need of adapted care and/or interventions. Differently, the Frailty Index acts as measure of the organism capacity to accumulate deficits. It tells us how many clinical conditions are present and concur at exhausting reserves. Thus, the Frailty Index seems to act as an objective marker of deficits accumulation.

2.2 Electronic Frailty Index (eFI)

Recently there has been some research efforts on developing an electronic frailty index (eFI) that can be automatically populated from routinely collected data contained within the primary care EHR. In the work of [12] a study was performed using anonymized primary care electronic health record data contained in massive databases (ResearchOne, The Health Improvement network (THIN)). The eligible patients were aged 65–95 years and had permanently registered at the practice. Using a scoring system, the patients were categorized into four categories of frailty:

1. **Fit (eFI score 0 - 0.12)** – People who have no or few long-term conditions that are usually well controlled. This group would mainly be independent in day to day living activities.
2. **Mild frailty (eFI score 0.13 – 0.24)** – People who are slowing up in older age and may need help with personal activities of daily living such as finances, shopping, transportation.
3. **Moderate Frailty (eFI score 0.25 – 0.36**) – People who have difficulties with outdoor activities and may have mobility problems or require help with activities such as washing and dressing.
4. **Severe Frailty (eFI score > 0.36)** – People who are often dependent for personal cares and have a range of long-term conditions/multimorbidity. Some of this group may be medically

The deficits that were identified by the authors and were used to generate the eFI score are listed in Table 2. These deficits are not homogeneous, as some refer to symptoms (i.e. dizziness, memory/cognitive problems), some are connected to disability (hearing/visual impairment) and some state the appearance of a disease (osteoporosis, Parkinson). Moreover there is significant overlap among them. For example, Fragility fracture and Osteoporosis are thwo different things, but quite similar in many cases. A more graphical way to organize the deficits is depicted in Figure 1.

Activity limitation	Memory and cognitive problems
Anemia and hematinic deficiency	Mobility and transfer problems
Arthritis	Osteoporosis
Atrial fibrillation	Parkinsonism and tremor
Cerebrovascular disease	Peptic ulcer
Chronic kidney disease	Peripheral vascular disease
Diabetes	Polypharmacy

Dizziness	Requirement for care
Dyspnea	Respiratory disease
Falls	Skin ulcer
Foot problems	Sleep disturbance
Fragility fracture	Social vulnerability
Hearing impairment	Thyroid disease
Heart failure	Urinary incontinence
Heart valve disease	Urinary system disease
Housebound	Visual impairment
Hypertension	Weight loss and anorexia
Hypotension/syncope	
Ischemic heart disease	

Table 2: List of the 36 deficits used in the eFI

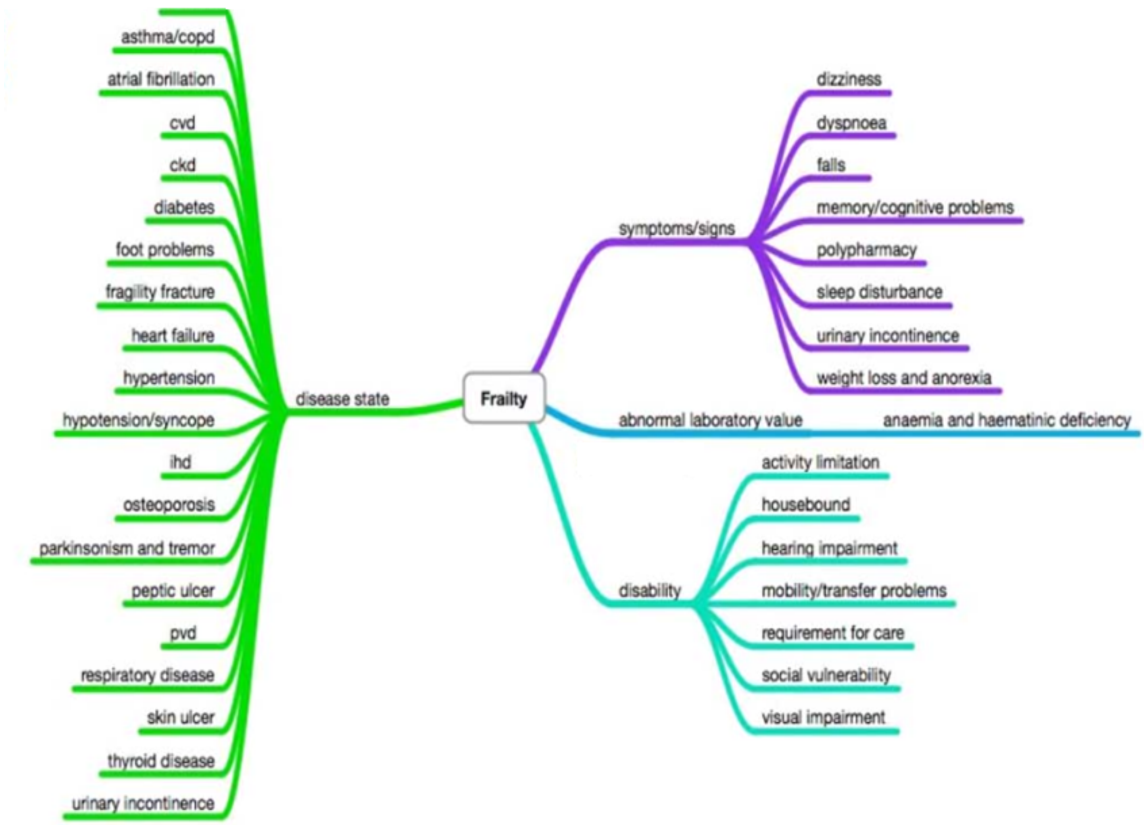

Figure 1: Map of the 36 deficits used in the eFI

The eFI has robust predictive validity and good discrimination for nursing home admission, hospitalization and mortality. These outcomes are of particular importance for older people and health and social care systems internationally, and the predictive validity and discrimination characteristics of the eFI across all three outcomes adds considerable weight to the clinical utility of the tool in terms of individual and population health planning.

2.3 Definition of Frailty Indices in FrailSafe

The clinical partners of the consortium analyzed the state of the art of frailty definitions as part of deliverable D2.1, and defined new frailty indices which are going to be used in FrailSafe. These indices aim to define the loss of reserve, independently of frailty status as this is defined by Fried's criteria, in order to render clinical results measurable.

On the other hand, FrailSafe Database contains variables at different time points from

- Clinical Evaluation
- Follow up assessment
- FrailSafe system metrics

In this scope, a new combined index (**Combined Frailty index: CoFI**), that will express frailty status relevant to the study's measurements, will be created by adding up two other frailty indices derived from the study, the **Clinical Frailty Index (ClFI)**, corresponding to the results of the clinical evaluation, and the **Technical Frailty Index (TFI)**, corresponding to the metrics derived from the FrailSafe system devices.

Each time a programmed clinical evaluation is effectuated, a ClFI score will be calculated, which will be composed by several items that correspond to various aspects of frailty, as they are described by the clinical evaluation sub-questionnaires. Similarly, a TFI will be calculated for each FrailSafe system installation, practically, for each FrailSafe home visit. Finally, a combined FI, by adding up ClFI and TFI will be calculated. A summary of these indices in shown in Table 3 and more can be found in deliverable D2.1.

Table 3: Frailty Indices definition

ClFI (Clinical Frailty Index): score corresponding to the findings of the clinical evaluation in a time-spot
TFI (Technical Frailty Index): accumulated score derived from the FrailSafe system metrics during certain time intervals of observation
CoFI (Combined Frailty Index): combined Clinical and Technical frailty score

Additionally we plan to link and associate FrailSafe parameters to the eFI parameters at a higher level of abstraction at this stage so that we will be able to evaluate our population using their scoring system. This will help us exploit the results of the eFI which were based on 900k+ health records and validated in large international studies. This way, we will be able to strengthen the statistic viability of our study, whilst at the same time being able to assess the added value of the FrailSafe system to our participants. The fusion of information from the huge in participant numbers eFI database to the very specialized, qualitative and highly personalized FrailSafe system will undoubtedly yield results of great significance as to the establishment of informative indicators and biomarkers for frailty.

3 SIGNAL PROCESSING AND DATA MINING TECHNIQUES FOR EXTRACTING FRAILTY RELATED INDICATORS

3.1 Analysis using tensors

The traditional approach to data representation utilizes a matrix structure, with observations in the rows and features in the columns. Although this model is appropriate for many datasets, it is not always a natural representation because it assumes the existence of a single target variable and lacks a means of modeling dependencies between other features. Additionally, such a structure assumes that observed variables are scalar quantities by definition. This assumption may not be valid in certain domains where higher-order features predominate, or in domains which have strong spatial, temporal or spatiotemporal components (e.g. ECG signals).

Traditionally, these problems have been solved by reducing the features to scalars and fitting the dataset to a matrix structure. However, as well as potentially losing information, this strategy also employs a questionable approach from a philosophical standpoint: attempting to fit the data to an imprecise model rather than attempting to accurately model the existing structure of the data. Finally, while it may be possible to model dependencies between features by repeating the methodology multiple times, each with a different target variable, this yields suboptimal performance and may not be computationally feasible when real-time performance is required or when the dataset is very large.

To address these issues, we propose to model such datasets using *tensors*, which are generalizations of matrices corresponding to multidimensional arrays. To formulate the use of tensors, we first need to establish some basic notation. The *order* of a tensor is the number of its dimensions. So a 3rd order tensor is a three-dimensional array, like the one shown in Figure 2.

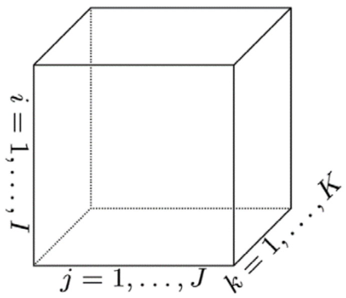

Figure 2: Third order tensor (source [13]).

Now if someone is dealing with *N* dimensions, the corresponding tensor will be an N^{th} order tensor. It is very common that tensors are treated either as sets of *fibers* or as sets of *slices.* A fiber is the higher order analogue of matrix row and column and it is defined by fixing all but one indices of a tensor. A slice is a two dimensional part of a tensor and it is defined by fixing all but two indices of a tensor. Figure 3 depicts shows all possible slices and fibers of a 3rd order tensor.

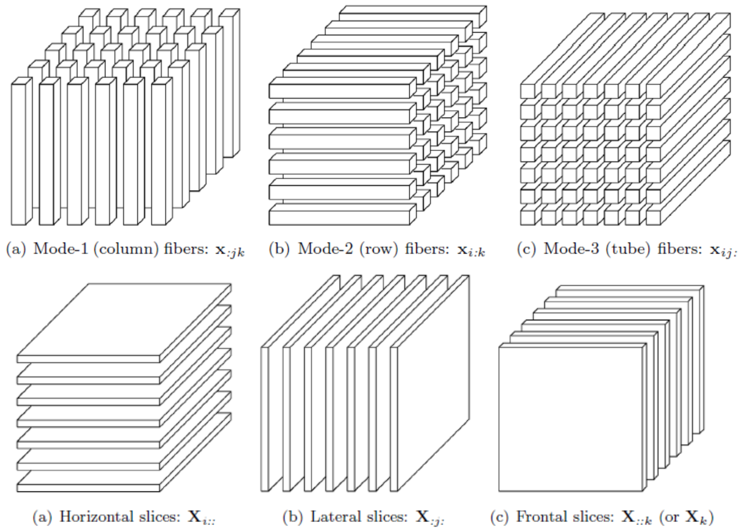

Figure 3: Slices and fibers of a third order tensor (source [13]).

3.1.1 Tensor decomposition

Having defined the structure of a tensor, one must go further to examine the value of representing their data with a tensor. This is where tensor decompositions enter to manipulate even the most pretentious sets of datasets with high dimensionality. Tensor decompositions, which are an extend of matrix decompositions coming from linear algebra, have a wide range of application including data mining, information retrieval, neuroscience, signal processing and many other problems. Their success lies on their ability to capture multi-linear and multi-aspect structures of high-dimensional datasets. The two most widely used tensor decomposition models are PARAFAC and Tucker. PARAFAC, as well as Tucker decomposition are the higher dimensional analogous of the widely known methods *Principal Component Analysis (PCA)* and *Singular Value Decomposition (SVD)*.

Singular value decomposition (SVD) is a unique matrix factorization by which a $m\times n$ matrix is decomposed into two projection matrices and a core matrix, as follows:

$$\mathbf{A} = \mathbf{U}\times\mathbf{\Sigma}\times\mathbf{V}^{\mathbf{T}}$$

where $\mathbf{A}$ is an $m\times n$ matrix, $\mathbf{U}$ is an $m\times r$ column-orthonormal projection matrix, $\mathbf{V}$ is an $n\times r$ column-orthonormal projection matrix, and $\mathbf{\Sigma}$ is a diagonal $r\times r$ *core matrix*, where r is the rank of the projection.

Singular value decomposition has a wide variety of applications: for example, truncation of the SVD coefficients provides an optimal low-rank approximation (i.e. minimizes the Frobenius norm). This indicates a close relationship between SVD and Principal Component Analysis (PCA). SVD is also used to discover the rank of a matrix, find the pseudoinverse, and solve least squares minimization problems. Additionally, the solution to SVD may be used in an unsupervised summarization technique known as Latent Semantic Analysis (LSA) [14]. In this technique, $\mathbf{A}$ is treated as a term-document matrix. Here, singular value decomposition automatically derives a user-specified number of latent *concepts* from the given terms which form a basis for the rows and columns of the matrix. The projection matrices $\mathbf{U}$ and $\mathbf{V}$ then contain term-to-concept and document-to-concept similarities, respectively. Thus, SVD can be used to provide simple yet powerful automatic data summarization. This technique may be naturally viewed as a form of co-clustering, in which the rows and columns of a matrix cluster to the same space. An alternative graphical interpretation exists, in which clusters represent shared “waypoints” through which edges pass between vertices. Use of the eigendecomposition or SVD is also common in a graphical context, where it is known as *spectral graph theory*; here a common technique is to cluster on the eigenvector corresponding to the second smallest eigenvalue of the Laplacian matrix, thereby partitioning vertices along edges which are likely to be minimal cuts. This technique is known as *Fiedler retrieval*. It is also possible to project new query vectors into the space defined by the SVD, known as *folding in*; this enables recommendation as the query projects to the same space as both the rows and columns and can be assessed using a distance metric.

The natural extension of singular value decomposition to tensors is known as *high-order singular value decomposition*, or HOSVD. This decomposition, in turn, is a special case of the *Tucker decomposition*, which is capable of concurrent data co-clustering across every mode of a tensor. Formally, let X be a tensor of order *R*; i.e. $X \in \Re^{d_1 \times d_2 \times \ldots \times d_R}$. We may then define the Tucker decomposition as the following factorization into a *core tensor* and a product of *r* projection matrices:

$$X = \mathcal{G} \times \mathbf{U}_1 \times \mathbf{U}_2 \times \ldots \times \mathbf{U}_r$$

The matrices of the decomposition are known as factor matrices and can be thought of as the principal components in each dimension (just like in the case of PCA). The core tensor’s entries show the level of interaction between the components. Depending on the number of columns of each factor matrix, the Tucker decomposition can be a compressed version of the original tensor X. A schematic representation for a 3rd order tensor is shown in Figure .

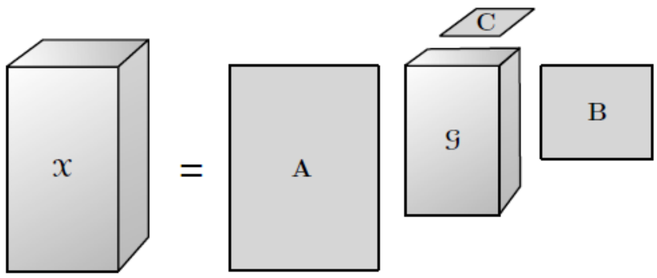

Figure 4: Tucker decomposition of a third order tensor (source [13]).

Note that while either the core tensor must be diagonal or the projection matrices must be column-orthonormal, the Tucker decomposition does not guarantee that both conditions are simultaneously true. When the projection matrices are unitary, the factorization is called high-order singular value decomposition.

When used as a data summarization technique, the Tucker decomposition exhibits similar behavior to singular value decomposition. Specifically, the core tensor’s elements represent the strengths of the discovered concepts (in terms of variance captured), while the projection matrices each represent the strength of the individual term-to-concept relationships on their corresponding modes.

Another method that has been popular in the field of tensor analysis is the one called both parallel factor analysis (PARAFAC) and canonical decomposition (CANDECOMP) due to simultaneous discovery of the method in 1970 by Harshman [15] and Carrol and Chang [16]. We will refer here to the method as PARAFAC.

PARAFAC [15] is a generalization of PCA and forms the basis of our tensor analysis approach. Given a user-specified number of concepts c, PARAFAC decomposes an order-r tensor X into a columnwise sum of the tensor product of r projection matrices, denoted $\mathbf{U}^{(1)} \ldots \mathbf{U}^{(r)}$. Formally, we define the decomposition as follows:

$$\mathcal{A} = \sum_{\mathrm{i}=1}^{\mathrm{c}} \lambda_{\mathrm{i}} \mathbf{U}_{:,i}^{(1)} \otimes \mathbf{U}_{:,i}^{(2)} \otimes \ldots \otimes \mathbf{U}_{:,i}^{(r)}$$

PARAFAC decomposition factorizes a tensor into a sum of rank-one tensors. A rank-one tensor is a3[rd] order tensor X that can be written as the outer product of *3* vectors. To that end, a PARAFAC decomposition of a 3[rd] order tensor can be schematically represented as shown in Figure 5.

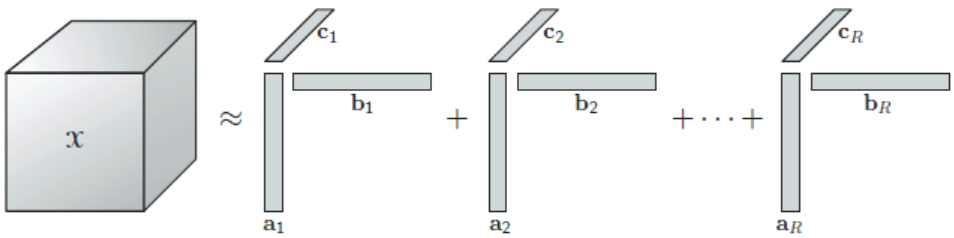

Figure 5: PARAFAC decomposition of a third order tensor (source [13]).

It can easily be seen that the PARAFAC decomposition is defined by using the fibers representation of a tensor. Applications of the aforementioned decomposition on data mining problems in the first place included discussion detanglement in online chat rooms [17] and automatic conversation detection in e-mail over time [18]. Applications on neuroscience have proven that PARAFAC successfully focuses on features of interest when it comes to analyze functional connectivity in the brain, revealing crucial information about changes in correlation strength between different locations (electrodes).

Just like in the PARAFAC case, Tucker decomposition has been used in the field of data mining for the problem of discussion detanglement in online chat rooms, as well as identifying handwritten digits, analyzing web site click-through data and several other applications.

A wide range of algorithms has been explored for the implementation of the aforementioned tensor decomposition models. The basic idea of most of them include using either the Alternating Least Square (ALS) [19] method or the Higher Order SVD method (HOSVD). During the past years many variations of these algorithms have been implemented, all of them aiming at reducing computational time and minimizing the size of the resources needed to compute the decomposition. Of course, the tradeoff between computational time/space and accuracy is a bit of a challenge, but that is always the case when dealing with large datasets.

Moving a step beyond the original use of the tensor decompositions, there has been a great effort to exploit the tensors' structure and tools considering several data mining problems including clustering, feature extraction and classification. A feature extraction and classification problem from a tensor decomposition point of view can be defined as follows:

Consider a set of K training samples (a set of arrays formulated as slices within a tensor) corresponding to C classes and a set of test data (test slices of a tensor). The challenge is to find appropriate labels for the test data. The latter can be performed in the following steps:

- *Find a set of basis matrices (just like in the Tucker model) and corresponding features for the training data.*
- *Perform feature extraction for test samples using the basis factors from the previous step.*
- *Perform classification by comparing the test features with the training features.*

One can easily understand that the above problem is an extension of the Tucker decomposition model, considering the factor matrices as the basis matrices, and the core tensor as the feature representation. The compressed core tensor is of much lower dimension than the original data, making it a fruitful option for dealing with the classification problem using as little resources and time as possible. The above method was proposed by Phan and Cichocki in [20] and was tested for handwritten digits, BCI motor imagery and image classification. A simplified scheme is shown below.

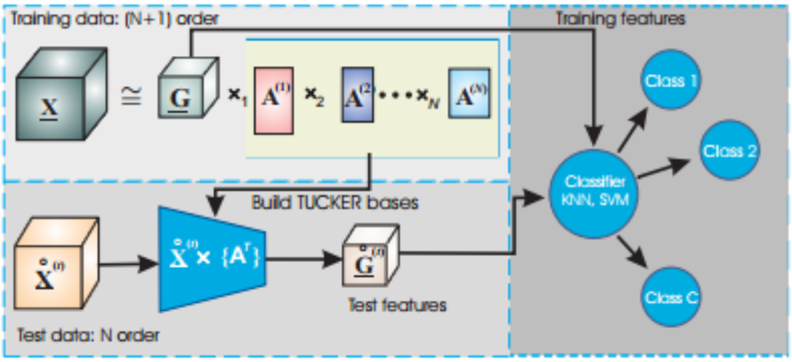

Figure 4: Classification diagram based on TUCKER decomposition (source [20]).

Another interesting approach on the co-clustering problem was introduced in the work of [21], where the idea was to use PARAFAC decomposition with sparse latent factors in order to extract tri-clusters from the original data. The uniqueness of the decomposition along with the sparsity constraint impose that a large number of possibly overlapping co-clusters will arise.

From the methods and applications mentioned above, it is clear that tensors and their decompositions are a precious tool for a wide range of fields, and provide the opportunity to extract hidden information of high dimensional datasets using state-of-the-art algorithms applicable to most common systems. The data analysis procedure requires a great amount of effort to be accomplished, and heading towards the right tools is the key to a successful result.

3.1.2 Applying tensors to FrailSafe data

Considering the multimodal nature of the data collected through the FrailSafe system development, the analysis should be headed towards representing sensory, physiological and device-related data with tensors. More specifically, each data source can be thought of as a dimension (a mode) in an *N*-dimensional tensor (where *N* is the number of the different kind of data sources). For each different sample of the data collected (a sample can be a set of data coming from a specific time course), there is a distinct tensor created, which in fact belongs to a set of *K* training tensors as mentioned above in the feature extraction & classification example. For each of these samples we suppose that the knowledge about the frailty condition exists (aka the class to which a training tensor belongs to). Performing a Tucker decomposition for the whole set of *K* tensors will conclude to having a set of features for each sample, living in the core tensor of each decomposition. The latter simplified representation is computationally expensive, and for that reason all sample tensors can be concatenated into a single tensor preferably vectorized, whose core tensor will include all features from all samples. The vectorized core tensor will include all possible features.

Up until that point, it is hopefully clear that even though the size of the data volume is important, manipulating the data as vectors guarantees that the data analysis at least offline is viable. Moving further the analysis, after extracting the features from the tensor training set classification should be performed. The choice of the classifier must be done after experimenting with the input data and features extracted. Hopefully feature ranking and feature selection will conclude to a stable set of features, which will account for the frailty indicator. A tensor test set of will be given as input to the classifier in order for each sample to be assigned to a frailty-dependent class.

As shown above, the use of tensor representation and tensor decompositions for the data provided during the FrailSafe system development, is highly recommended for the purpose of frailty indicator extraction. The computational cost of the analysis is low, making it an attractive solution for a wide range of data analysis problems, including FrailSafe.

Looking at the analysis procedure from a different perspective, data fusion of multiple-source data can be formulated as a coupled matrix and tensor factorization (CMTF) problem. Coupled factorization techniques arose from the need to jointly analyze heterogeneous datasets, meaning datasets with different order, in the form of matrices or higher-order tensors. The formulation of factorizing a 3rd order tensor $X \in \mathbb{R}^{I\times J\times K}$ coupled with a matrix $Y \in \mathbb{R}^{I\times M}$ is the following

$$f(A,B,C,V) = \|X - [\![A,B,C]\!]\|^2 + \|Y - AV^T\|^2 ,$$

where the factorization of X and Y is performed through the minimization of the above equation which fits a CANDECOMP/PARAFAC (CP) model to X and factorizes Y so that the factor matrix corresponding the common mode i.e. $A \in \mathbb{R}^{I\times R}$ is the same. Factor matrices $B \in \mathbb{R}^{J\times R}$ and $C \in \mathbb{R}^{K\times R}$ correspond to the second and third modes of the tensor. $V \in \mathbb{R}^{M\times R}$ is the factor matrix that corresponds to the second mode of Y. The above formulation is ideal for revealing underlying structures in joint datasets when all factors are shared across datasets. But it fails to capture factors only in the presence of both shared and unshared components. To overcome this issue, the problem is reformulated as

$$f(\lambda,\Sigma,A,B,C,V) = \|X - [\![\lambda;A,B,C]\!]\|^2 + \|Y - A\Sigma V^T\|^2 + \gamma\|\lambda\|_1 + \gamma\|\sigma\|_1 ,$$

where the columns of factor matrices have unit norm and $\boldsymbol{\lambda} \in \mathbb{R}^{R\times1}$ and $\boldsymbol{\sigma} \in \mathbb{R}^{R\times1}$ are the weights of rank-1 components in the third-order tensor and the matrix respectively. $\boldsymbol{\Sigma} \in \mathbb{R}^{R\times R}$ is a diagonal matrix with entries of $\boldsymbol{\sigma}$ in the diagonal. Finally, γ>0 is a penalty parameter. Adding the constraints of weighting and penalty parameters, unshared components will have weights equal or close to zero in one of the datasets. The latter approach can be further explored in [22], [23], [24]. A schematic representation of the reformulated problem is shown below.

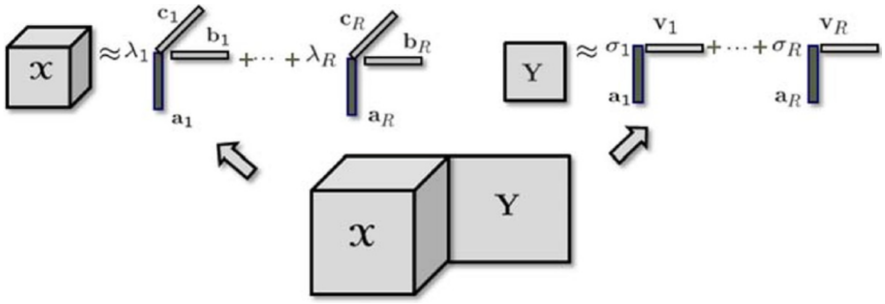

Figure 5: Constrained CMTF model (source [24]).

Experimenting with a variety of optimization algorithms, a solution to the problem of structure-revealing CMTF model can be found in order to assess robust results through the joint analysis of complementary sources. Especially in the case of FrailSafe where sensory data can be represented as a high-dimensional tensor while a matrix can contain location-related (or physiological) data, fusion through a coupled matrix and tensor factorization scheme could yield extraordinarily satisfying results. Tensor-based aided methods such as the one mentioned above for the purpose of knowledge discovery through multiple-sources datasets are considered a great option in the context of FrailSafe system development. Properly extracted underlying structures through the joint analysis can bring into light hidden frailty-related components, which in turn will be given as an input to the Virtual Patient Model as well as the monitoring system of FrailSafe. Picking robust methods in order to set up a valuable human-oriented system is the cornerstone of our success.

3.2 Mining of multi-level association rules

Towards discovering associations between frailty, and physiological or behavioral patterns, a preliminary work has been made in association rules. The aim is to discover a novel way of mining multi-level association rules, in a distributed environment, from multiple heterogeneous data sources.

Nowadays, as a result of cheap storage and data availability, the volume of data is huge and is expanding rapidly and very often there is the need to discover useful and interesting knowledge from quite different data sources. While methodologies and solutions exist to mine rules effectively in a single node environment, these methodologies fail when data volume expands beyond a threshold. On the other hand, distributed systems and platforms have appeared to present an alternative processing model, capable of handling effectively massive loads of data. Despite their immense capabilities, these systems lack established methodologies in order to fully exploit their resources.

Our goal is to take the positive features from both models and combine them into a unified model, capable of handling massive data volume and performing established knowledge discovery methodologies (association rule mining) on them. Our focus is twofold:

- Combination of association rules and concept hierarchies to be able to mine multi-level rules from unified and augmented data sources
- Exploitation of the processing power and capabilities of distributed systems to effectively handle the increased data volume

To accomplish our objective, we augment input data, based on concept hierarchies (which are adjusted on the problem at hand), to produce a unified and augmented data file. This file is sent to a distributed processing system (Hadoop framework stack) to generate large frequent itemsets effectively, based on the procedure proposed by the Apriori algorithm. Multi-level association rules are then produced from these itemsets and pruned based on optimization parameters, in order to keep only those that are interesting.

3.2.1 Architectural Model

The model that is being developed is based on a multi-tier architecture. The different tiers are presented in Figure 6, where Configuration tier is marked in blue, Processing tier is marked in green, Output tier is marked in red and finally Control tier is marked in yellow. The packages out of which the tiers are consisted, are explained below.

1. **Configuration and input**: This package is responsible for every configuration that is needed at the system setup phase. System variables, hierarchy files, input and output folders, ssh connections as well as checking access rights for files and folders, where needed. Additionally, tests on the structure and format of input files (data files, hierarchy files, configuration file) are carried out, in order to ensure a smooth system initialization. By design, direct user input for system variables and hierarchies will be available, along with automated importing from xml files.
2. **Data augmentation:** Data augmentation is used to implement the idea of multi-level concept hierarchies. This procedure is performed on all input files, based on the hierarchy files provided at the configuration stage, in order to unify all data entries into a single augmented file. Information for the hierarchy structure is stored along with the actual data (that are considered to be at leaf level), following a bottom up approach.
3. **Distributed processing:** The distributed system that will be used is the Hadoop framework and its software stack. Hadoop provides many tools for distributed processing (Mahout, MapReduce, Spark), but they require coordination during the intermediate phases and the many and time consuming iterations of the Apriori large frequent itemset generation procedure. Moreover, the actual communication, as well as input and output, with the Hadoop environment has specific issues that must be considered, especially in the case of remote ssh-based communication.
4. **Rule generation and pruning:** Large frequent itemset generation is the most costly part of the Apriori algorithm, after which the rule production follows. Rules are based on the generated itemsets and are associated with several metrics and thresholds, in order to estimate their value. The degree of interestingness determines whether a rule will be pruned or not. The various thresholds are determined during the configuration phase and metric values are exported along with the rules.
5. **Output:** All the interesting rules that have remained, along with their metrics, are stored into a report (in xml format) and the report is exported as system output.
6. **Control:** The coordination between the various stages and phases of our model is done by this package. Due to the differences in input and output

data structure and of course features, there must be interfaces between the modules, in order to ensure their efficient operation.

The flow diagram of the overall system is summarized in Figure 7.

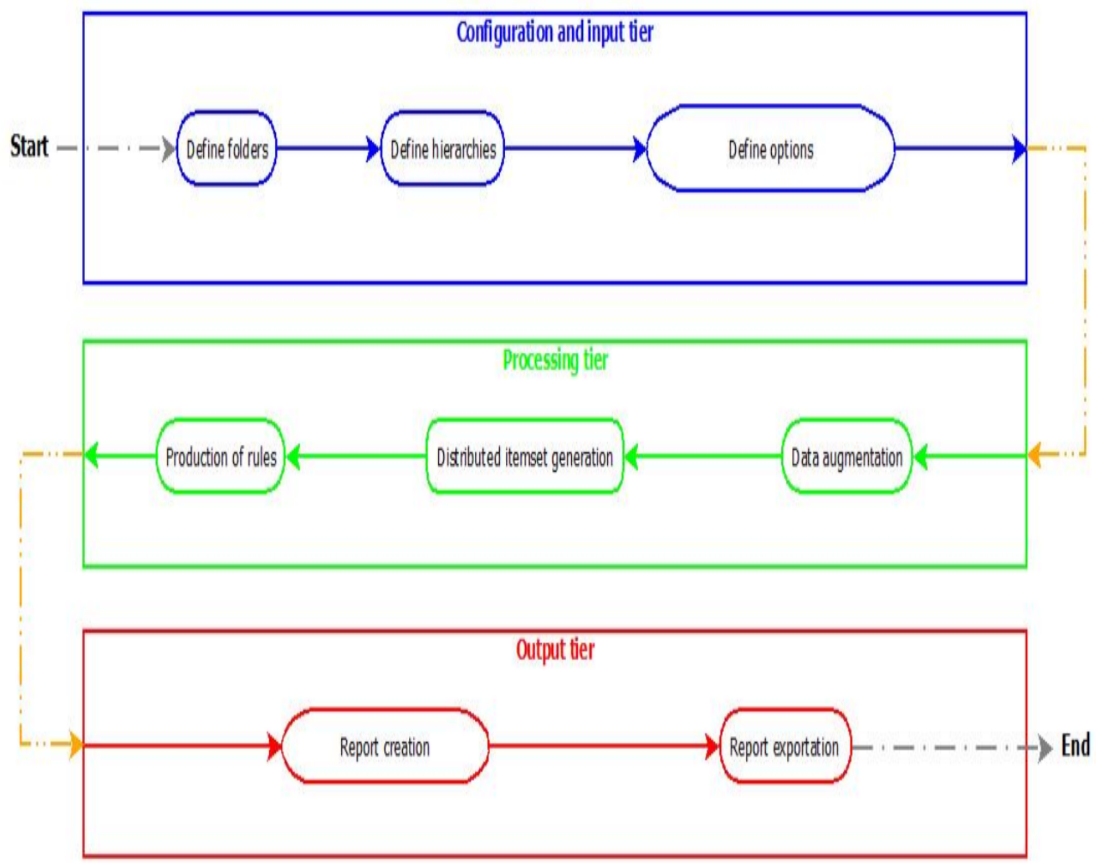

Figure 6: Multi-tier model

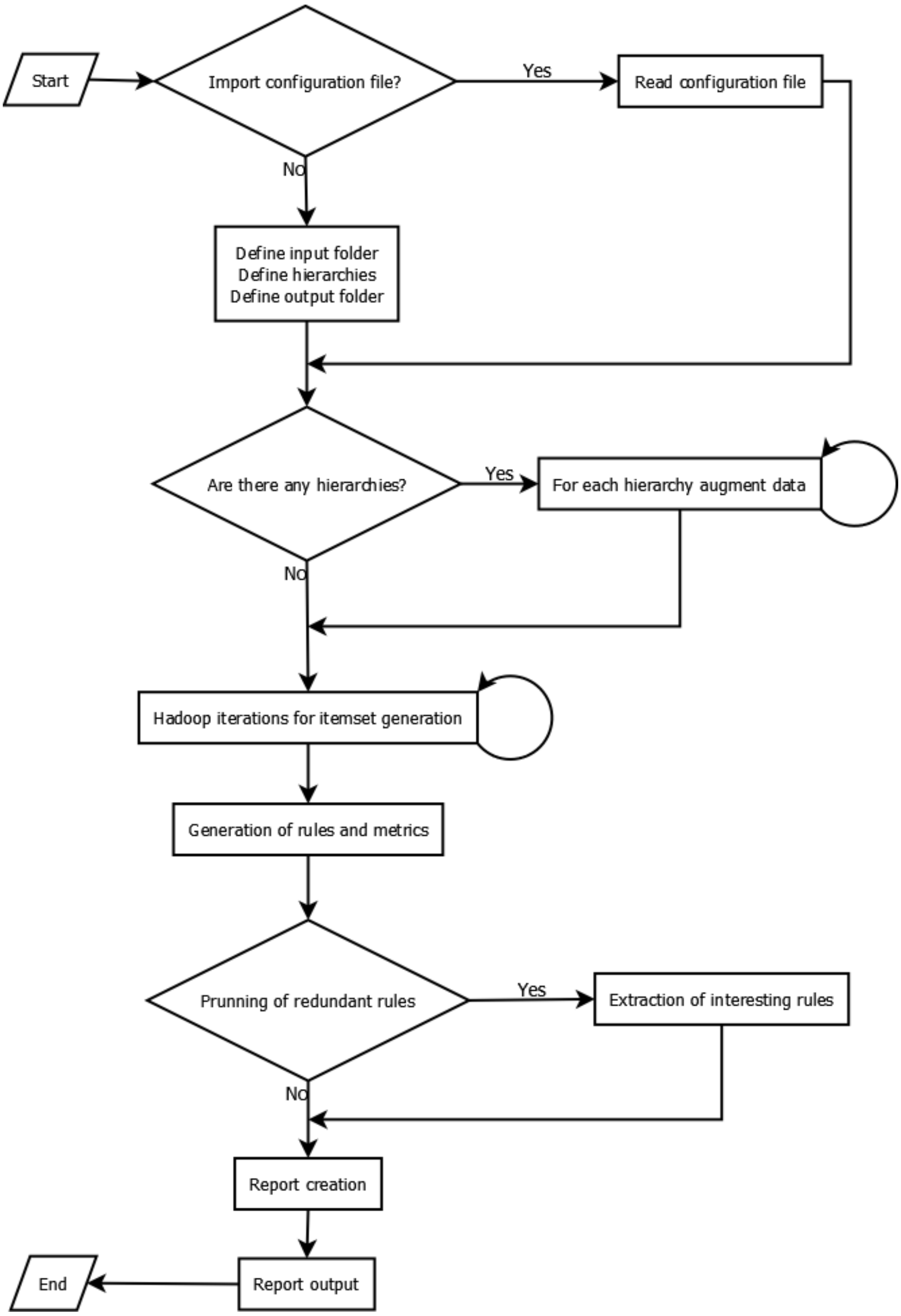

Figure 7: System control flow diagram

3.2.2 Current and future work

As a preliminary work, the state of the art in multi-level association rule methods has been explored. Several recent works ([25], [26], [27]) are studying the creation and the organization of a multi-level concept hierarchy, along with the data augmentation procedure. The implementation of association mining algorithms such as the Apriori in distributed architectures has been in ([28], [29], [30]). Motivated by the related work, the basic architecture and packet layout have been designed and the source codes for a number of functions (retrieving hierarchy data and creating the corresponding xml files, reading hierarchy data from xml files, creation and reading of the xml configuration file) have been developed in java. Furthermore existing algorithms in Mahout, from the Hadoop framework stack, are examined in order to generate the frequent itemsets. As a secondary choice, an Apriori implementation based on the main map-reduce Hadoop framework stack could be used. The proposed model is going to be tested using FrailSafe's multimodal data:

- Questionnaire data
- Data collected from the mobile phone
- Data collected from the vest
- Data study and preprocessing are required to determine which attributes will be used and how

In the following period our efforts will be focused on developing, implementing and testing our model. The next steps can be summarized as follows:

- Comprehensive study and analysis of available data and their structure, to determine the precise augmentation process
- Study and testing of the distributed system features and their interaction with the rest of the system
- Completion of the distinct software packages, as regards to their basic functionality
- Implementation of the basic software functionality (for the entire system), in order to run tests and configure the system for maximum performance (few hierarchies vs many hierarchies, distributed vs non-distributed implementation, handling multiple data sources)
- Creation of complete test cases (after the implementation of the basic system functionality), that cover everything from system input (data sources) to system output (exported report)
- Expansion of the basic software features, based on the designed multi-tier model, to their full extent
- System testing, using actual data as input and evaluation of its usability, performance and efficiency
- External user reviews (lab members that are unfamiliar with the system) for feedback on the various features
- Final system evaluation, using the complete data set and comparing the various configurations and their results

4 DATA FUSION

There are two main approaches for fusing data from different sensor units/dimensions: feature-level fusion and decision-level fusion. In feature-level fusion, which is commonly used to exploit the dependencies across dimensions, the data are fused directly after feature extraction. Feature vectors from each dimension/sensor unit are fused and events are classified by one global classifier. On the other hand, in decision-level fusion, events are classified for each dimension/sensor unit by its local classifier and the results from these local classifiers are later fused in the decision layer.

Analysis of multi-sensor data is very complex and difficult to summarize with a small number of variables extracted from the multi-dimensional signals. As a result, analysis is usually accompanied by extraction of high dimensional feature vectors from data. The dimensionality is further increased in feature-level fusion approaches aiming to exploit the information across dimensions/sensor units, where already high dimensional feature vectors from several sensor units are combined to a single large feature vector. The problem of high dimensionality coupled with limited number of samples, usually available in practice, makes the analysis of multidimensional signals a challenging task.

Thus we propose a new decision-level scheme to deal with the problem of high dimensionality in conjunction with small number of samples. The proposed scheme combines information from all sensor units in order to train a single classification model and thus is sensor-independent. The decision-level fusion scheme keeps the dimensionality quite low, while the incorporation of a global training model allows the use of more training samples (by combining all sensor units).

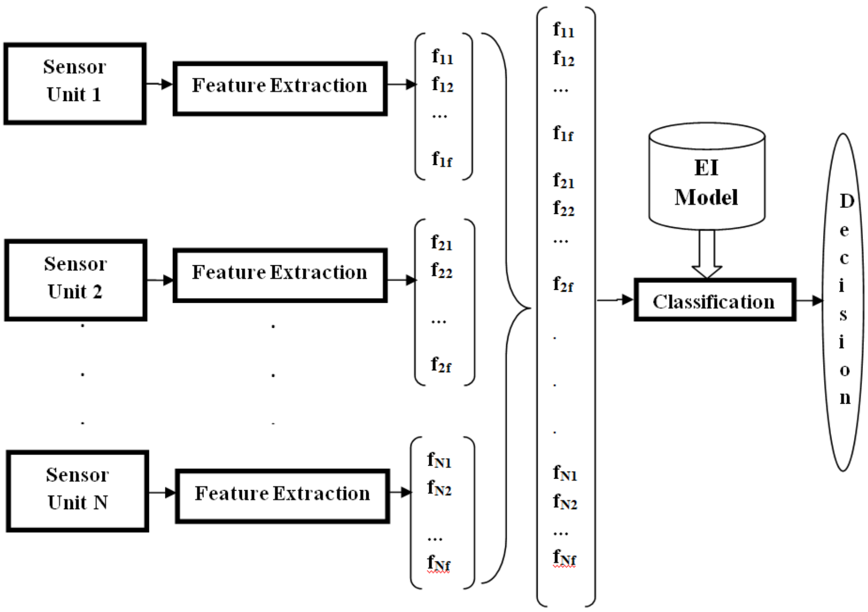

Figure 8: Feature-level fusion scheme.

Feature Level fusion

In the feature-level scheme, each one of the available sensor units from each frame is processed in parallel by the feature extraction algorithm. The estimated feature vectors from each sensor unit are concatenated into a single feature vector. This 'super' feature vector is used as a representative signature for the corresponding frame. Therefore, the training set is a data matrix $M \times N \times f$, where M is the number of frames in the training set, N is the number of sensor units and f the number of features extracted from each sensor unit. The feature level scheme is illustrated in Figure 8. Although such a scheme exploits the information from all dimensions of the data, it leads to a feature vector of high dimensionality imposing the need either for feature selection before classification, or the availability of a large number of training samples.

Decision-level fusion with local training models

In the decision-level fusion with local (sensor dependent) training models, a separate classification model is built for each sensor unit. Each one of the available sensor units is processed in parallel by the feature extraction algorithm and the estimated feature vectors are used to form N training sets, one for each sensor unit. The data matrix of each training set is $M \times f$ here. For each frame, N decisions are made by each one of the N local classifiers. A final decision is made by combining the N output class labels using a fusion rule, such as majority voting. The decision-level with local training models fusion scheme is illustrated in Figure 9. In decision-level fusion schemes the dimensionality of the feature vector is smaller than in feature-level fusion schemes. However, this scheme uses training samples only of the corresponding sensor unit.

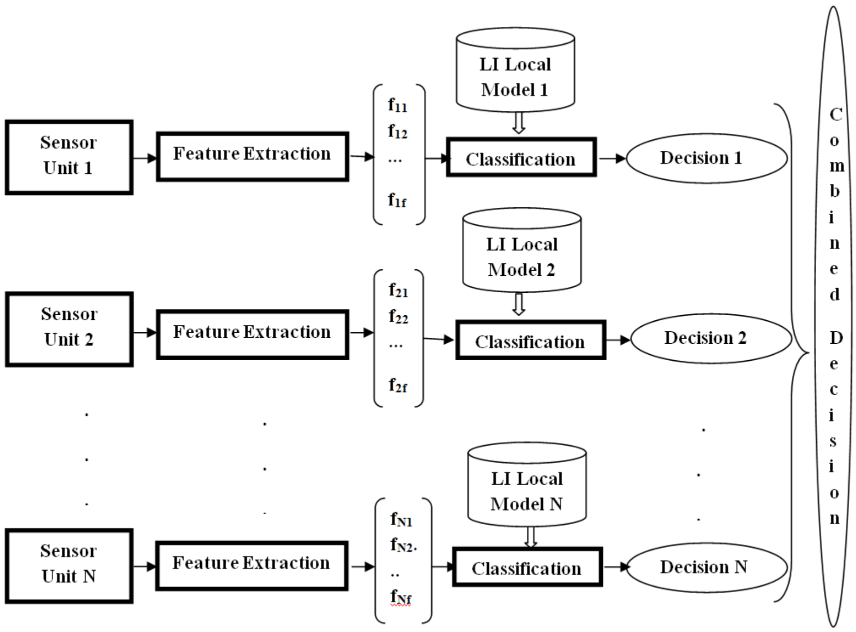

Figure 9: Decision-level with local (sensor dependent) training models fusion scheme.

Decision-level with global training model

In the decision-level with global (sensor independent) training model fusion scheme, a common classification model is used for the feature vectors extracted from the different sensor units. The data matrix of the training set is now $N \times M \times f$ and is constructed by merging all training sets from the decision-level with local models fusion scheme. In this scheme the number of training samples is larger since each data frame appears in the training set N times, one time for each one of the available sensor units. During the test phase, for each frame, N decisions are made by feeding the signature from each sensor unit to the global classification model. A final decision is made at a score level by combining the N output class labels using the same fusion rule (majority voting) as before. The decision-level with global training model scheme is illustrated in Figure 10.

Although this scheme is less specific, it handles better both the high dimensionality and the problem of small number of training instances.

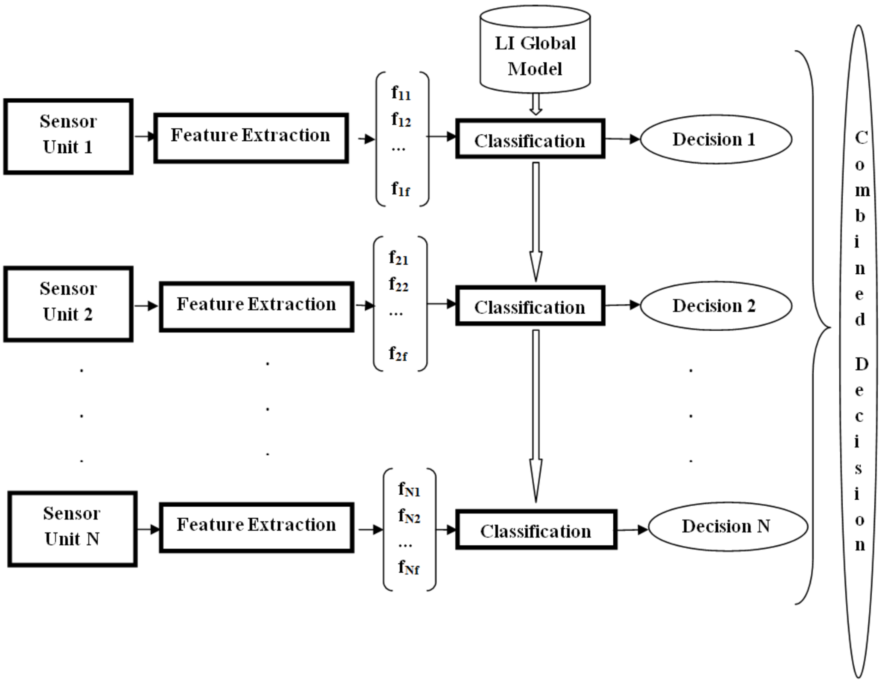

Figure 10: Decision-level with global (sensor independent) training model fusion scheme.

Our preliminary work in data fusion schemes were tested on different datasets and showed promising results [31]. In the next period we intend to use the older person data collected through FrailSafe, in order to validate our schemes.

REFERENCES

[1] A. Clegg, J. Young, S. Iliffe, M. O. Rikkert and K. Rockwood, "Frailty in elderly people," *The Lancet,* vol. 381, pp. 752-762, 2013.

[2] J. Subra, S. Gillette-Guyonnet, M. Cesari, S. Oustric, B. Vellas and others, "The integration of frailty into clinical practice: preliminary results from the Gerontopole," *The journal of nutrition, health \& aging,* vol. 16, pp. 714-720, 2012.

[3] L. P. Fried, J. Walston and others, "Frailty and failure to thrive," *Principles of geriatric medicine and gerontology,* vol. 5, pp. 1487-502, 2003.

[4] L. P. Fried, C. M. Tangen, J. Walston, A. B. Newman, C. Hirsch, J. Gottdiener, T. Seeman, R. Tracy, W. J. Kop, G. Burke and others, "Frailty in older adults evidence for a phenotype," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences,* vol. 56, pp. M146--M157, 2001.

[5] K. Rockwood, X. Song, C. MacKnight, H. Bergman, D. B. Hogan, I. McDowell and A. Mitnitski, "A global clinical measure of fitness and frailty in elderly people," *Canadian Medical Association Journal,* vol. 173, pp. 489-495, 2005.

[6] K. E. Ensrud, S. K. Ewing, B. C. Taylor, H. A. Fink, P. M. Cawthon, K. L. Stone, T. A. Hillier, J. A. Cauley, M. C. Hochberg, N. Rodondi and others, "Comparison of 2 frailty indexes for prediction of falls, disability, fractures, and death in older women," *Archives of internal medicine,* vol. 168, pp. 382-389, 2008.

[7] R. J. J. Gobbens, A. L. M. van Assen and Marcel, K. G. Luijkx, M. T. Wijnen-Sponselee and J. M. G. A. Schols, "The Tilburg frailty indicator: psychometric properties," *Journal of the American Medical Directors Association,* vol. 11, pp. 344-355, 2010.

[8] A. S. Buchman, P. A. Boyle, R. S. Wilson, Y. Tang and D. A. Bennett, "Frailty is associated with incident Alzheimer’s disease and cognitive decline in the elderly," *Psychosomatic medicine,* vol. 69, pp. 483-489, 2007.

[9] M. Cesari, G. Gambassi, G. van Kan and Abellan and B. Vellas, "The frailty phenotype and the frailty index: different instruments for different purposes," *Age and ageing,* vol. 43, pp. 10-12, 2014.

[10] L. P. Fried, L. Ferrucci, J. Darer, J. D. Williamson and G. Anderson, "Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences,* vol. 59, pp. M255--M263, 2004.

[11] G. A. Kuchel, "Frailty, allostatic load, and the future of predictive gerontology," *Journal of the American Geriatrics Society,* vol. 57, pp. 1704-1706, 2009.

[12] A. Clegg, C. Bates, J. Young, R. Ryan, L. Nichols, E. A. Teale, M. A. Mohammed, J. Parry and T. Marshall, "Development and validation of an electronic frailty index using routine primary care electronic health record data," *Age and ageing,* vol. 45, pp. 353-360, 2016.

[13] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review,* vol. 51, pp. 455-500, 2009.

[14] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer and R. Harshman, "Indexing by latent semantic analysis," *Journal of the American society for information science,* vol. 41, p. 391, 1990.

[15] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multi-modal factor analysis," *UCLA Working Papers in Phonetics,* 1970.

[16] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition," *Psychometrika,* vol. 35, pp. 283-319, 1970.

[17] E. Acar, S. A. Çamtepe, M. S. Krishnamoorthy and B. Yener, "Modeling and multiway analysis of chatroom tensors," in *International Conference on Intelligence and Security Informatics*, 2005.

[18] B. W. Bader, M. W. Berry and M. Browne, "Discussion tracking in Enron email using PARAFAC," in *Survey of Text Mining II*, Springer, 2008, pp. 147-163.

[19] R. Sands and F. W. Young, "Component models for three-way data: An alternating least squares algorithm with optimal scaling features," *Psychometrika,* vol. 45, pp. 39-67, 1980.

[20] A. H. Phan and A. Cichocki, "Tensor decompositions for feature extraction and classification of high dimensional datasets," *Nonlinear theory and its applications, IEICE,* vol. 1, pp. 37-68, 2010.

[21] E. E. Papalexakis, N. D. Sidiropoulos and R. Bro, "From k-means to higher-way co-clustering: Multilinear decomposition with sparse latent factors," *IEEE transactions on signal processing,* vol. 61, pp. 493-506, 2013.

[22] E. Acar, M. A. Rasmussen, F. Savorani, T. Næs and R. Bro, "Understanding data fusion within the framework of coupled matrix and tensor factorizations," *Chemometrics and Intelligent Laboratory Systems,* vol. 129, pp. 53-63, 2013.

[23] E. Acar, M. Nilsson and M. Saunders, "A flexible modeling framework for coupled matrix and tensor factorizations," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, 2014.

[24] E. Acar, E. E. Papalexakis, G. Gürdeniz, M. A. Rasmussen, A. J. Lawaetz, M. Nilsson and R. Bro, "Structure-revealing data fusion," *BMC bioinformatics,* vol. 15, p. 239, 2014.

[25] K. Koperski and J. Han, "Discovery of spatial association rules in geographic information databases," in *International Symposium on Spatial Databases*, 1995.

[26] R. Srikant and R. Agrawal, Mining generalized association rules, IBM Research Division, 1995.

[27] T.-L. B. Tseng, C.-C. Huang, K. Fraser and H.-W. Ting, "Rough set based rule induction in decision making using credible classification and preference from medical application perspective," *Computer methods and programs in biomedicine,* vol. 127, pp. 273-289, 2016.

[28] J. Li, P. Roy, S. U. Khan, L. Wang and Y. Bai, "Data mining using clouds: An experimental implementation of apriori over mapreduce," in *Proc. 12th International Conference on Scalable Computing and Communications (ScalCom'13)*, 2012.

[29] J. Woo, "Apriori-Map/Reduce Algorithm," in *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA)*, 2012.

[30] X. Y. Yang, Z. Liu and Y. Fu, "MapReduce as a programming model for association rules algorithm on Hadoop," in *Information Sciences and Interaction Sciences (ICIS), 2010 3rd International Conference on*, 2010.

[31] E. Pippa, E. I. Zacharaki, M. Koutroumanidis and V. Megalooikonomou, "Data fusion for paroxysmal events' classification from EEG," *Journal of Neuroscience Methods,* vol. 275, pp. 55-65, 2017.