Using the Choquet Integral for the Determination of the Anxiety Degree

Ion Chițescu

University Politehnica of Bucharest

Mădălina Giurgescu Manea (✉ madalina_giurgescu@yahoo.com)

University of Pitesti

Titi Paraschiv

Military Technical Academy

Research Article

Keywords: Choquet Integral, Anxiety Degree, psychologists

Posted Date: November 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1027357/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Using the Choquet Integral for the Determination of the Anxiety Degree

Ion Chițescu¹, Mădălina Giurgescu Manea^{2,*}, and Titi Paraschiv³

¹Department of Mathematics, University Politehnica of Bucharest, Splaiul Independenței 313, Bucharest, 060042, Romania

²Department of Mathematics, University of Pitesti, Târgul din Vale 1, Pitesti, 110040, Romania

³Military Technical Academy of Bucharest, George Coșbuc 39-49, Bucharest, 050141, Romania

*madalina.giurgescu@yahoo.com

ABSTRACT

This paper introduces a mathematical model describing how the EEG type waves are processed in order to characterize the level of anxiety. The electroencephalogram (EEG) is a recording of the electrical activity of the brain. The main frequencies of the human EEG waves are: Delta, Theta, Alpha (Low Alpha and High Alpha), Beta (Low Beta and High Beta), Gamma. Psychologists' studies show that there is an interactive relationship between anxiety and two factors in the Big Five theory, namely, extraversion and neuroticism. The specialists in psychology state that the anxiety is characterized by LowAlpha, HighAlpha, LowBeta and HighBeta waves. In this regard, we developed a mathematical model through which EEG waves are processed in order to determine the level of anxiety. Our main idea is to use the Choquet integral with respect to a suitable monotone measure in order to characterize the level of anxiety. This measure was obtained using the measurements of the values of EEG waves made on 70 subjects and the corresponding levels of anxiety (established by psychologists) of these subjects. In order to verify our mathematical model (aggregation tool) we used it to determine the level of anxiety of 10 other subjects, comparing our results with the results provided by psychologists (the comparison validated our results).

1 Introduction

In this paper, we describe a mathematical model used for processing the level of EEG type waves in order to characterize the level of anxiety, which represents changes in the values of the personality characteristics of the BigFive model (see⁶).

The electroencephalogram (EEG) is a recording of the electrical activity of the brain. The recorded waveforms reflect the cortical electrical activity (see¹).

The main frequencies of the human EEG waves are:

- Delta: is a high amplitude brain wave with a frequency of oscillation between 0.5 and 4 Hz. Delta waves, like other brain waves, are recorded with an electroencephalogram (EEG) and are usually associated with the deep stage 3 of NREM sleep, also known as slow-wave sleep (SWS) and aid in characterizing the depth of sleep (see¹).

- Theta: has a frequency of 4 to 7 Hz and is classified as "slow" activity. Cortical theta is observed frequently at young children. At older children and adults, it tends to appear during meditative, drowsy, hypnotic or sleeping states, but not during the deepest stages of sleep. Several types of brain pathology can give rise to abnormally strong or persistent cortical theta waves (see¹).

- Alpha: has a frequency between 8 and 13 Hz. It is usually best seen in the posterior regions of the head on each side, being higher in amplitude on the dominant side. It appears when closing the eyes and relaxing and disappears when opening the eyes or alerting by any mechanism (thinking, calculating). It is the major rhythm seen at normal relaxed adults. It is present during most of life especially after the thirteenth year. The overall alpha band may be divided into lower and higher alphas. Lower Alpha (6–10 Hz) is a response to any type of task and is topographically spread over almost all electrodes. Higher Alpha, restricted to parietooccipital areas, is found during visually presented stimulations (see¹).

- Beta: beta activity is "fast" activity. It has a frequency between 13 and 30 Hz. It is usually seen on both sides in symmetrical distribution and is most evident frontally. It is accentuated by sedative-hypnotic drugs, especially the benzodiazepines and the barbiturates. It may be absent or reduced in areas of cortical damage. It is generally regarded as a normal rhythm. It is the dominant rhythm at patients who are alert or anxious or have their eyes open. Beta waves can be split into three sections: Low Beta Waves (12.5–16 Hz); Beta Waves (16.5–20 Hz); and High Beta Waves (20.5–28 Hz) (see¹).

- Gamma: is a pattern of neural oscillation at humans with a frequency between 30 and 140 Hz, the 40-Hz point being of particular interest. Gamma rhythms are correlated with large scale brain network activity and cognitive phenomena such as

working memory, attention and perceptual grouping and can be increased in amplitude via meditation or neurostimulation. Altered gamma activity has been observed in many mood and cognitive disorders (see¹).

In psychological trait theory, the Big Five personality traits, also known as the five-factor model (FFM), is a suggested taxonomy, or grouping, for personality traits, developed from the 1980s onwards. The Big Five theory is based therefore on semantic associations between words and not on neuropsychological experiments. This theory uses descriptors of common language and suggests five broad dimensions commonly used to describe the human personality and psyche (see⁶). The theory identifies five factors (see⁶):

- openness to experience (inventive or curious vs. consistent or cautious)
- conscientiousness (efficient or organized vs. extravagant or careless)
- extraversion (outgoing/energetic vs. solitary/reserved)
- agreeableness (friendly or compassionate vs. challenging or callous)
- neuroticism (sensitive or nervous vs. resilient or confident)

Anxiety is an emotional state characterized by a feeling of insecurity, disorder, diffuse (see²). It is a state of fear, agitation, insecurity and nervousness. Symptoms of anxiety include high blood pressure, helplessness, maladaptation, sadness and anxiety, manifested in the body by palpitations (fast heartbeat), trembling, sweating in the palms and insomnia (see²).

It is estimated that approximately about 6.5% of the world's population suffers or has suffered from medically diagnosed anxiety, but there are many more who suffer from the annoying symptoms of stress or anxiety (see²).

Psychologists' studies show that there is an interactive relationship between anxiety and two factors in the Big Five theory, namely, extraversion and neuroticism. More exactly, anxiety has a positive relationship with neuroticism and a negative relationship with extraversion (see⁶).

The specialists in psychology state that the anxiety is characterized by LowAlpha, HighAlpha, LowBeta and HighBeta waves (see¹).

In this regard, we have developed a mathematical model through which EEG waves are processed in order to determine the level of anxiety.

Before exhibiting our mathematical model described in the present paper, we shall make some bibliographical comments, pertaining to some aspects of the EEG theory. A very good and documented material, dedicated to the history of the encephalography, beginning with Hans Berger, can be found in¹⁴.

The basic material concerning EEG considered by us, when writing this paper, is contained in the monographs^{1,2} and in the glossary³. In⁴, the role of Alpha oscillations in cognitive psychomotor, psycho-emotional and physiological aspects of human life is discussed. In⁶ one lays stress upon the interpretation of the fact that brain oscillation and empirical evidence link Delta oscillation with reward motivation and Alpha oscillation with anxiety. In⁸ an interesting use of the Fourier transform in the study of the encephalograms is proposed. In⁹, the Delta band (1-3.5 Hz) of the EEG oscillatory activity is studied, linked to a broad variety of perceptual and cognitive operations. In¹¹ one lays- stress upon the fact that, up to now, reactive, as opposite to proactive behaviour, during social interaction, have not been investigated in relation to other kinds of social behaviour. A virtual interaction model is proposed. In¹³ one explains that the association between neuroticism and anxiety may be additionally explained by transdiagnostic factors.

It is known that, as a mathematical procedure, one frequently uses non-linear integrals as fusion instruments (see^{5,7,12,16} and¹⁷). We shall take this into account, our target being to determine an instrument of aggregation which uses the data resulted from the measurements, thus obtaining conclusions on the subjects in regard to anxiety. The main idea of the present paper is that this instrument should be the Choquet integral with respect to some virtual monotone measure, whose existence we accept and whose values are computed, according to some devices described in the sequel.

In writing this paper we have worked in collaboration with the Institute of Studies, Research, Development and Innovation of Titu Maiorescu University in Bucharest, as well as with specialists of the Military Technical Academy in Bucharest.

The measurements of the values of EEG waves were performed on 70 subjects. In order to carry out the measurements, a NeuroSky device, with two sensors, of which one active, was used. We only used EEG wave values for Low Alpha, High Alpha, Low Beta, and High Beta waves.

The input data used were specific values of EEG waves, as well as classification data of the anxiety level, provided by the Psychology Research Institute. The classification given by the psychologists to the subjects (measured individuals), is represented by grades, from 0 to 6 (meaning that 0 is the least anxious and 6 the most anxious). For EEG measurements, see¹ and³.

To calculate the values, we have created a C++ program (see¹⁵). The source code is written in C++ in the CodeBlocks development medium, 17.12 version on Windows 10 operating system, combined with GNU GCC Compiler in MinGW distribution, 6.3 version. For the matrix operations the Eigen library, version 3.3 was used.

We determined an aggregation tool through which we drew conclusions regarding the level of anxiety for 10 new subjects, comparing the results from our model with the results provided by psychologists, the results obtained after the comparison

being very good.

So, the target was to determine an instrument of aggregation, by using the data resulted from the measurements and using the conclusions on the subjects in regard to anxiety, as well as to use this instrument to draw conclusions in regard to the level of anxiety of some subjects who have undergone new measurements. As we said, the instrument proposed by us is the aforementioned Choquet integral.

We think this procedure is new, considering (among others) the fact that generally, statistical methods are used, or psychological tests (see, e.g.¹⁰).

2 Preliminary facts

Throughout the paper, the positive integer numbers will be $\mathbb{N} = \{0, 1, 2, \dots\}$, the real numbers will be $\mathbb{R}$ and the positive real numbers will be $\mathbb{R}_+ = \{x \in \mathbb{R} | x \geq 0\}$. As usual, $\bar{\mathbb{R}}_+ = \mathbb{R}_+ \cup \{\infty\}$.

For any set T , the Boolean of T is $P(T) = \{A | A \subset T\}$. A measurable space is a couple (T, τ) , where T is a non-empty set and $\tau \subset P(T)$ is a σ -algebra.

If (T, τ) is a measurable space, a monotone measure (or a fuzzy measure) is a function $\mu : \tau \rightarrow \mathbb{R}_+$ having the properties:

- i) $\mu(\emptyset) = 0$
- ii) $\mu(A) \leq \mu(B)$ for any A, B in τ such that $A \subset B$

Now, let us consider a measurable space (T, τ) , a monotone measure $\mu : \tau \rightarrow \mathbb{R}_+$ and a positive τ -measurable function $f : T \rightarrow \mathbb{R}_+$.

For any $a \in \mathbb{R}_+$, we consider the inferior level set $F_a = \{t \in T | f(t) \geq a\} = f^{-1}([a, \infty)) \in \tau$. Because $F_b \subset F_a$, whenever $0 \leq a < b < \infty$, it is seen that the function $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, given via $\phi(a) \stackrel{\text{def}}{=} \mu(F_a)$, is decreasing. Considering the Lebesgue measurable sets of $\mathbb{R}_+$ which we denote by $\mathcal{L}$ and the Lebesgue measure on $\mathcal{L}$ which we denote by $L : \mathcal{L} \rightarrow \mathbb{R}_+$, we can compute $\int \phi dL$, because ϕ is $\mathcal{L}$ -measurable. For the sake of concreteness and respecting the traditional notations, we shall write $\int_0^\infty \mu(F_a) da$ instead of $\int \phi dL$.

Definition 1. The Choquet integral of the function f with respect to the measure μ is the element $\int_0^\infty \mu(F_a) da \in \bar{\mathbb{R}}_+$. We shall write

$$(C) \int f d\mu = \int_0^\infty \mu(F_a) da$$

We shall say that f is Choquet integrable with respect to μ in case $(C) \int f d\mu < \infty$.

The Choquet integral is a generalization of the abstract Lebesgue integral. Namely, in case μ is a classic measure (i.e. μ is σ -additive), we can see that the Choquet integral $(C) \int f d\mu$ coincides with the abstract Lebesgue integral $\int f d\mu$.

We have special formulae for the computation of the Choquet integral in case the function f is simple, i.e. f has the form

$$f = \sum_{i=1}^n a_i \phi_{A_i}$$

where $a_i \in \mathbb{R}_+$, $A_i \in \tau$ are mutually disjoint and $\cup_{i=1}^n A_i = T$. Here ϕ_A is the characteristic (indicator) function of the set $A \subset T$.

Namely, for such f , one can put in order the numbers a_i such that $a_1 \leq a_2 \leq \dots \leq a_n$. Considering that f is in this situation, one has the formula

$$(C) \int f d\mu = \sum_{i=1}^n (a_i - a_{i-1}) \mu(A_i \cup A_{i+1} \cup \dots \cup A_n) \quad (1)$$

with the convention $a_0 \stackrel{\text{def}}{=} 0$.

Caution: The reordering of the values a_i is unique in case a_i are distinct and all A_i are nonempty. Otherwise, different reorderings can occur, but the value of $(C) \int f d\mu$ given by the formula from above does not depend upon these different reorderings.

For Generalized Measure and Integration Theory, see^{7, 12, 16} and¹⁷. Some special computing devices for the Choquet integral appear in⁵.

3 Using the Choquet integral to solve the inverse problem of information fusion

3.1 Linearization Formula for The Computation of The Finite Discrete Choquet Integral

We shall be concerned with the case when T is finite, $T = \{x_1, x_2, \dots, x_n\}, n \geq 1$. Hence, for the function $f : T \rightarrow \mathbb{R}_+$, there exists a permutation $\sigma : T \rightarrow T$ (we write $\sigma(x_i) = x_i^*$ for any $i = 1, 2, \dots, n$) such that $f(x_1^*) \leq f(x_2^*) \leq \dots \leq f(x_n^*)$.

In case the values $f(x_1), f(x_2), \dots, f(x_n)$ are distinct, such σ is unique.

Working for the measurable space $(T, \tau) = (T, P(T))$ and for a monotone measure $\mu : P(T) \rightarrow \mathbb{R}_+$, the formula 1 from Preliminary Facts gives

$$(C) \int f d\mu = \sum_{i=1}^n (f(x_i^*) - f(x_{i-1}^*)) \mu(\{x_i^*, x_{i+1}^*, \dots, x_n^*\}) \quad (2)$$

with the convention $f(x_0^*) = 0$.

Let us write $P^*(T) = \{A \subset T \mid A \neq \emptyset\} = P(T) \setminus \{\emptyset\}$ (hence $P^*(T)$ has $2^n - 1$ sets). For any $E \in P^*(T)$, define

$$a_E \stackrel{\text{def}}{=} \min_{x_p \in E} f(x_p) - \max_{x_q \in T \setminus E} f(x_q)$$

with the convention $\max_{x_q \in \emptyset} f(x_q) = 0$ (in the case $E = T$).

Next, define

$$b_E = \begin{cases} a_E, & \text{if } a_E \geq 0, \\ 0, & \text{if } a_E < 0 \end{cases}$$

It is seen that $b_E = a_E^+$.

The next theorem gives a linear formula, with respect to μ , for the computation of $(C) \int f d\mu$ (no need to order the values of f). This formula is crucial for solving the inverse problem of information fusion. We give the proof in order to make the paper self-contained.

Theorem 1. *One has the formula*

$$(C) \int f d\mu = \sum_{E \in P^*(T)} b_E \mu(E) \quad (3)$$

Proof. The result is obvious in case $n = 1$. From now on, we shall work for $n > 1$.

Define

$$M \stackrel{\text{def}}{=} \sum_{E \in P^*(T)} b_E \mu(E)$$

Notice that we have the inequality

$$(C) \int f d\mu \leq M \quad (4)$$

Indeed, any term of the form

$$(f(x_i^*) - f(x_{i-1}^*)) \mu(\{x_i^*, x_{i+1}^*, \dots, x_n^*\})$$

in the sum giving $(C) \int f d\mu$ (see (2)) is of the form $b_E \mu(E)$ where $E = \{x_i^*, x_{i+1}^*, \dots, x_n^*\}$, because $f(x_i^*) = \min_{x_p \in E} f(x_p) \geq f(x_{i-1}^*) = \max_{x_q \in T \setminus E} f(x_q)$ and $b_E = f(x_i^*) - f(x_{i-1}^*)$ (including the situation $E = T$ too).

In order to prove that (4) is actually an equality, one must prove that the possible new terms (i.e. terms in the sum giving M which do not appear in the sum giving $(C) \int f d\mu$ in (2)) are null.

So, take $E \in P^*(T)$ and let us examine $b_E \mu(E)$.

The case $E = T$ is trivial, because in this case $b_E \mu(E) = f(x_1^*) \mu(T)$ and this term appears in the sum giving $(C) \int f d\mu$ in (2). From now on, consider $E \neq T$ and compute a_E . There are three possibilities: $a_E = 0$, $a_E < 0$ and $a_E > 0$. In case $a_E \leq 0$, one has $b_E = 0$, hence $b_E \mu(E) = 0$, irrespective of the fact the $b_E \mu(E)$ appears or not in the sum giving $(C) \int f d\mu$ in (2) (in case $a_E < 0$ it is clear that E cannot be of the form $E = \{x_i^*, x_{i+1}^*, \dots, x_n^*\}$, hence $b_E \mu(E)$ does not appear in the sum giving $(C) \int f d\mu$ in (2)). Finally, in case $a_E > 0$, hence $b_E = a_E$, we shall see that the term $b_E \mu(E) = a_E \mu(E)$ appears in the sum giving $(C) \int f d\mu$ in (2), due to the fact that E must have the form $E = \{x_i^*, x_{i+1}^*, \dots, x_n^*\}$ for some $2 \leq i \leq n$.

Indeed, because $a_E > 0$, one has $f(x_q) < f(x_p)$ for any $x_q \in T \setminus E$ and any $x_p \in E$. For any rearrangement (permutation) of $x_1, x_2, \dots, x_n$ in the form $x_1^*, x_2^*, \dots, x_n^*$ such that $f(x_1^*) \leq f(x_2^*) \leq \dots \leq f(x_n^*)$, we have

$$\max_{x_q \in T \setminus E} f(x_q) < \min_{x_p \in E} f(x_p)$$

Consequently, we must find $i \in \{2, \dots, n\}$ such that $\max_{x_q \in T \setminus E} f(x_q) = f(x_{i-1}^*)$ and $\min_{x_p \in E} f(x_p) = f(x_i^*)$, which means that $E = \{x_i^*, x_{i+1}^*, \dots, x_n^*\}$. $\square$

3.2 Canonical Enumeration of $P^*(t)$

In order to have complete and concrete recipes for the computation of $(C) \int f d\mu$, we shall introduce the canonical enumeration of $P^*(T) = \{E_1, E_2, \dots, E_{2^n-1}\}$, the order of the enumeration $E_1, E_2, \dots, E_{2^n-1}$ being explained in the sequel.

Any number $j = 1, 2, \dots, 2^n - 1$ can be uniquely written in the binary form

$$j = \overline{j_n j_{n-1} \dots j_1} = j_1 + 2j_2 + 2^2 j_3 + \dots + 2^{n-1} j_n$$

where all $j_i \in \{0, 1\}$ and at least one $j_i \neq 0$.

The natural order of the numbers j coincides with the lexicographic order of the representative complexes $(j_n, j_{n-1}, \dots, j_1)$.

For instance, if $n = 3$, hence $2^n - 1 = 7$, each $j = 1, 2, \dots, 7$ has the form $j = \overline{j_3 j_2 j_1} = j_1 + 2j_2 + 2^2 j_3$: $j = 1 = \overline{001}$ (with $j_1 = 1; j_2 = 0; j_3 = 0$); $j = 2 = \overline{010} < j = 3 = \overline{011} < j = 4 = \overline{100} < j = 5 = \overline{101} < j = 6 = \overline{110} < j = 7 = \overline{111}$.

We shall enumerate $P^*(T)$ in lexicographic order: $E_1, E_2, \dots, E_7$, viewed as $E_1 = E_{\overline{001}}, E_2 = E_{\overline{010}}, \dots, E_7 = E_{\overline{111}} = E_{2^n-1}$. Practically, via this concrete exemplification, we defined the canonical enumeration of $P^*(T)$.

Consequently, the membership rule is the following:

$$x_i \in E_{\overline{j_n j_{n-1} \dots j_1}} \Leftrightarrow j_i = 1$$

(i.e. $E_{\overline{j_n j_{n-1} \dots j_1}} = \cup_{j_i=1} \{x_i\}$)

This membership rule generates, exactly $2^n - 1$ different sets.

For $n = 3$, one has: $E_1 = \{x_1\}, E_2 = \{x_2\}, E_3 = \{x_1, x_2\}, E_4 = \{x_3\}, E_5 = \{x_1, x_3\}, E_6 = \{x_2, x_3\}, E_7 = \{x_1, x_2, x_3\}$.

One can see that formula (3) in Theorem 1 can be written in the form

$$(C) \int f d\mu = \sum_{j=0}^{2^n-1} b_{E_j} \mu(E_j) \quad (5)$$

where any $j = 1, 2, \dots, 2^n - 1$ is written in the form $\overline{j_n j_{n-1} \dots j_1}$.

3.3 Solving The Inverse Problem of Information Fusion in This Case (Identification of The Monotone Measure Used To Generate The Aggregation Tool)

We shall consider the elements x_i of the set $T = \{x_1, x_2, \dots, x_n\}$ as the source of our information, e.g. any such x_i is a subject of our observation. Any observation of all subjects in T will be considered as a function $f : T \rightarrow \mathbb{R}$. We shall make l observations $f_1, f_2, \dots, f_l$, obtaining for any such observation $f_i : T \rightarrow \mathbb{R}$ the values $f_i(x_j), j = 1, 2, \dots, n$ and the conclusion (which is numerical) $y_i \in \mathbb{R}$. Namely, the values $f_i(x_j)$ and the values y_i are the input of the system (each y_i is a value of the fusion target). The idea is to use as aggregation tool the Choquet integral with respect to the some monotone measure μ . Acting in this manner, we obtain the data set as follows:

$$\begin{array}{cccccc} f_1(x_1) & f_1(x_2) & \cdots & f_1(x_n) & y_1 \\ f_2(x_1) & f_2(x_2) & \cdots & f_2(x_n) & y_2 \\ & & \cdots & & \\ f_l(x_1) & f_l(x_2) & \cdots & f_l(x_n) & y_l \end{array} \quad (6)$$

where we accept the existence of a monotone measure $\mu : P(T) \rightarrow \mathbb{R}_+$ such that

$$y_p = (C) \int f_p d\mu, p = 1, 2, \dots, l \quad (7)$$

Acting in the spirit of the inverse problem of information fusion, we shall consider that the unknown object is the measure μ .

To be more precise, we shall consider that the observed values $f_i(x_j)$ and the values y_i are known, but the measure μ is unknown (i.e. the values $\mu(E_j), j = 1, 2, \dots, 2^n - 1$ are unknown), in this case the measure being the output.

Taking into account formulae (5), (6), (7) it follows that the values $\mu(E_1), \mu(E_2), \dots, \mu(E_{2^n-1})$ must be solutions of the linear system (l equations with $2^n - 1$ unknowns).

$$\sum_{j=1}^{2^n-1} b_{pj} \mu(E_j) = y_p, p = 1, 2, \dots, l \quad (8)$$

Here, according to Theorem 1, we have, for any $p = 1, 2, \dots, l$ and any $j = 1, 2, \dots, 2^n - 1$, $b_{pj} = b_{E_j}^p$, i.e. $b_{pj} = a_{E_j}^p$ if $a_{E_j}^p \geq 0$ and $b_{pj} = 0$ if $a_{E_j}^p < 0$, where $a_{E_j}^p = \min_{x_t \in E_j} f_p(x_t) - \max_{x_t \notin E_j} f_p(x_t)$.

So, to solve our problem means to solve the system (8) and, generally speaking, this is a difficult task from computational point of view. It is preferable to solve (8) approximately, using the least squares method.

To this end, we consider the function $F : \mathbb{R}^{2^n-1} \rightarrow \mathbb{R}$, given via

$$F(t_1, t_2, \dots, t_{2^n-1}) = \sum_{p=1}^l \left(\sum_{j=1}^{2^n-1} b_{pj} t_j - y_p \right)^2$$

This function is infinitely differentiable and convex, being a sum of convex functions of the form $(t_1, t_2, \dots, t_{2^n-1}) \rightarrow (a_1 t_1 + a_2 t_2 + \dots + a_{2^n-1} t_{2^n-1} + b)^2$.

Hence, F has a minimum, attained at the point $(t_1^0, t_2^0, \dots, t_{2^n-1}^0)$ which is the solution of the system of equations

$$\frac{\partial F}{\partial t_k}(t_1^0, t_2^0, \dots, t_{2^n-1}^0) = 0, k = 1, 2, \dots, 2^n - 1$$

($2^n - 1$ equations with $2^n - 1$ unknowns t_j^0).

This system is in fact

$$\sum_{p=1}^l b_{pk} \left(\sum_{j=1}^{2^n-1} b_{pj} t_j^0 - y_p \right) = 0 \Leftrightarrow$$

$$\sum_{j=1}^{2^n-1} \left(\sum_{p=1}^l b_{pk} b_{pj} \right) t_j^0 = \sum_{p=1}^l b_{pk} y_p, k = 1, 2, \dots, 2^n - 1 \quad (9)$$

It is seen that, in case $(t_1^0, t_2^0, \dots, t_{2^n-1}^0) = (\mu(E_1), \mu(E_2), \dots, \mu(E_{2^n-1}))$ is an exact solution of (8), then $(t_1^0, t_2^0, \dots, t_{2^n-1}^0)$ is a solution of last system (9) too.

Otherwise, a solution $(t_1^0, t_2^0, \dots, t_{2^n-1}^0)$ of (9) is an approximate solution of (8). We shall consider the measure $\nu : P(T) \rightarrow \mathbb{R}_+$ given via $\nu(E_j) = t_j^0, j = 1, 2, \dots, 2^n - 1$ as the (approximate) solution of our problem.

The practical matrix solution of (9) is described in the sequel.

Consider the matrices

$$X \stackrel{\text{def}}{=} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ & & \vdots & \\ b_{l1} & b_{l2} & \cdots & b_{lm} \end{pmatrix}, \text{ of type } (l, m = 2^n - 1)$$

$$Y \stackrel{\text{def}}{=} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_l \end{pmatrix}, \text{ of type } (l, 1)$$

and the (approximate) matrix solution

$$A = \begin{pmatrix} t_1^0 \\ t_2^0 \\ \vdots \\ t_m^0 \end{pmatrix}, \text{ of type } (m = 2^n - 1, 1)$$

It is seen that (9) says that

$$X^T X A = X^T Y \quad (10)$$

In the (desirable) particular case when $X^T X$ is invertible, the solution A is given via

$$A = (X^T X)^{-1} X^T Y \quad (11)$$

Caution: It is advisable to work for $l \geq 2^n - 1$, i.e. one must perform a sufficiently large amount of observations. The reason for this restriction is given by the fact that

$$\text{rank}(X^T X) \leq \min(\text{rank}(X^T), \text{rank}(X)) = \text{rank}(X) \leq l$$

So, in case $l < 2^n - 1$, it follows that $X^T X$ is not invertible and the last formula for A (i.e. formula (11)) cannot be used.

Correction of the results:

The approximate solution v must be a monotone measure. The verification of this fact is the following: for any $E \in P^*(T)$, $E \neq T$, one must have

$$0 \leq v(E) \leq v(E \cup \{x_i\}), \text{ for any } x_i \in T - E$$

If this not in case (i.e. negative values $v(E)$ apper, or violation of the monotonicity occur), we modify v as follows (it is possible for some of subsequent steps to be unnecessary):

1. First, all (possible) stictly negative values $v(E)$ are replaced by the value 0 (hence all $v(G)$, with $G \subset E$, become 0 too).
2. Next, if there exist E, F in $P^*(T)$ such that $E \subset F$ and $v(E) > v(F)$, then we modify $v(E)$, replacing the value $v(E)$ with

$$\max\{v(G) | G \subset E \text{ and } v(G) \leq \min\{v(H) | E \subset H\}\}$$

(the set under max is not empty, containing $\emptyset$).

3. The procedure at 2. continues for the modified $v(E)$ until "wrong" pairs (E, F) as above do not appear any more.

Finally, we obtain the monotone measure μ which will be called "the good measure".

4 Determination of the anxiety degree

Main Idee of the Paper

We use the Choquet integral in order to determine the level of anxiety of a subject S .

To be more precise, we consider the total set $T = \{x_1, x_2, x_3, x_4\}$, where $x_1 = \text{LowAlpha}$, $x_2 = \text{HighAlpha}$, $x_3 = \text{LowBeta}$, $x_4 = \text{HighBeta}$. The function $f : T \rightarrow \mathbb{R}_+$ which will be integrated gives the input values $f(x_i)$, $i = 1, 2, 3, 4$ (results of the respective measurements for S). Namely, for each $i = 1, 2, 3, 4$, we extract from the CSV file the set of the values pertaining to i and we compute the average of these values, which is exactly $f(x_i)$. If $\mu : P(T) \rightarrow \mathbb{R}_+$ is the suitable measure proposed by us (its construction follows immediately). We accept that the level of anxiety of S is (approximately) equal to $(C) \int f d\mu$. This procedure is immediately exemplified in the sequel.

4.1 Creating The Instrument (Measure). The Inverse Problem

We made measurements for $l = 70$ subjects (we think, $l = 70$ is sufficient to furnish a credible measure). For each subject, the measurement values were provided in a CSV file. Using the obtained data we shall construct the $l = 70$ functions $f_1, f_2, \dots, f_{70}$. The $n = 4$ measured attributes are $x_1 = \text{LowAlpha}$, $x_2 = \text{HighAlpha}$, $x_3 = \text{LowBeta}$, $x_4 = \text{HighBeta}$. Thus, we obtained the table Tab. 1. with 70 rows and $4 + 1 = 5$ columns. Namely, for each of the 70 measurements (rows), we obtained the input values: $f_p(x_1), f_p(x_2), f_p(x_3), f_p(x_4)$ and the value y_p , $p = 1, 2, \dots, 70$. So, the fifth column contains the input values y_p , $p = 1, 2, \dots, 70$. We repeat the explanation concerning the values $f_p(x_i)$. For each subject p and each $i = 1, 2, 3, 4$ corresponding to one of the attributes LowAlpha, HighAlpha, LowBeta, HighBeta we extract from the CSV file the set of values pertaining to i and we compute the average of these values. The obtained result represents $f_p(x_i)$. The input values y_p are the classification given by the psychologists to the subjects (measured individuals), represented by grades, from 0 to 6 (meaning that 0 is the least anxious and 6 the most anxious).

As we have said, we decided to choose as fusion instrument the Choquet integral of the functions f_p , $p = 1, 2, \dots, 70$, with respect to a virtual(unknown) monotone measure μ which will be determined. The determined measure will be the output. So, for any $p = 1, 2, \dots, 70$ one has (similar to (7)):

$$y_p = (C) \int f_p d\mu \quad (12)$$

In the C++ program, we used a function to process the data from the CSV files, and to create a matrix. In order to save typographical space, we exhibit below only one row of the Table 1

Table 1. Sample data

Low Alpha	High Alpha	Low Beta	High Beta	Grade
41482.67	27173.99	18173.78	19565.17	6

Using the devices described in section 3, we determined the "measure" v . It has been necessary to (slightly) modify v (see correction of the Result, end of Section 3) and we obtained the good measure μ , whose values $\mu(E)$, arranged in the aforementioned lexicographic order, are the following:

$$\mu(E) = \begin{cases} 0 & \text{if } E = \emptyset \\ 0.0000208904, & \text{if } E = \{x_1\} \\ 0.0000104939, & \text{if } E = \{x_2\} \\ 0.000047098, & \text{if } E = \{x_1, x_2\} \\ 0.0000398943, & \text{if } E = \{x_3\} \\ 0.000230985, & \text{if } E = \{x_1, x_3\} \\ 0.000142402, & \text{if } E = \{x_2, x_3\} \\ 0.000230985, & \text{if } E = \{x_1, x_2, x_3\} \\ 0.0000143691, & \text{if } E = \{x_4\} \\ 0.0000357572, & \text{if } E = \{x_1, x_4\} \\ 0.0000143691, & \text{if } E = \{x_2, x_4\} \\ 0.000047098, & \text{if } E = \{x_1, x_2, x_4\} \\ 0.0000398943, & \text{if } E = \{x_3, x_4\} \\ 0.000230985, & \text{if } E = \{x_1, x_3, x_4\} \\ 0.000142402, & \text{if } E = \{x_2, x_3, x_4\} \\ 0.000230985, & \text{if } E = \{x_1, x_2, x_3, x_4\} \end{cases}$$

The measure was calculated using th C++ program.

4.2 Using The Obtained Instrument (Measure) to Determine The Anxiety Level of Other Subjects. The Direct Problem

We used the obtained measure μ to decide over the anxiety level of 10 new subjects. We made $l = 10$ measurements. The measurement values were provided in a CSV file. For each subject p , $p = 1, 2, \dots, 10$, we obtained the values $f_p(x_i)$, $i = 1, 2, 3, 4$ as previously. In this case the input data are $f_p(x_i)$ and the measure μ . We also have the conclusions of psychologists for these new topics with which we will compare our results. Namely, measuring these $l = 10$ new subjects, we obtained the values $f_p(x_1), f_p(x_2), f_p(x_3), f_p(x_4), p = 1, 2, \dots, 10$. See Table 2

Table 2. New subjects

No	Low Alpha	High Alpha	Low Beta	High Beta
1.	33738.85	26911.79	15911.23	15827.22
2.	11360.2	8108.433	9596.968	11737.35
3.	35170.35	23910.76	17211.32	12380.47
4.	36224.77	27315.48	16978.18	21143.96
5.	10658.68	7869.542	8243.346	7708.505
6.	33411.1	22809.05	21446.92	14769.47
7.	36568.82	23903.5	14545.97	19463.21
8.	33738.85	26911.79	15911.23	15827.22
9.	41482.67	27173.99	18173.78	19565.17
10.	51089.42	27487.33	19538.43	20055.94

The classification given by the psychologists to the subjects (measured individuals), is represented by grades, from 0 to 6 (meaning that 0 is the least anxious and 6 the most anxious). See Table 3

Table 3. Grades

No	Grades
1.	4
2.	2
3.	5
4.	5
5.	2
6.	5
7.	4
8.	4
9.	5
10.	5

Using the formula (12), we obtained the conclusions(the output data):

$$\begin{aligned}
 y_1 &= 4.33598 \approx 4, \\
 y_2 &= 2.28522 \approx 2, \\
 y_3 &= 4.52631 \approx 5, \\
 y_4 &= 4.59469 \approx 5, \\
 y_5 &= 1.95455 \approx 2, \\
 y_6 &= 5.23955 \approx 5, \\
 y_7 &= 4.06521 \approx 4, \\
 y_8 &= 4.33598 \approx 4, \\
 y_9 &= 4.92068 \approx 5, \\
 y_{10} &= 5.38052 \approx 5.
 \end{aligned}$$

As can be seen, the resulting conclusions are very close to the conclusions provided by psychologists.

This procedure can be continued for any new subject, given that the results obtained in this way are very good.

5 Conclusions

1. The studied level of anxiety represents changes in the values of the personality characteristics in the BigFive model, and its values were determined using EEG waves.
2. We have determined an aggregation model, using the data resulted from the measurements and using the conclusion on the subjects regarding the anxiety. The aggregation model consists in a monotone measure and the Choquet integration of the input data with respect to this measure.
3. This determined measure was successfully used to draw conclusions regarding the level of anxiety of new subjects measured with NeuroSky.
4. The created model was validated by comparing its values with the ones obtained via classical methods (psychological tests).

References

1. Electroencephalography, Basic Principles, Clinical Applications and Related Fields (E. Niedermeyer and F. Lopes da Silva, eds), 4th edition. Lippincott, Williams and Wilkins, Philadelphia, Pennsylvania, 1999.
2. Diagnostic and Statistical Manual of Mental Disorders. DSM-5. 5th edition. American Psychiatric Association, Washington DC, 2013

3. A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017. Corrected 2019. Clin.Neurophy.Pract, 4, 133, 2019.
4. O.M.Bazanova, D.Vernon: Interpreting EEG alpha activity. Neuroscience and Biobehavioral Reviews, vol. 44, pp 94–110, 2014.
5. I.Chişescu, A.Plăviţu: Computing Choquet integrals. Fuzzy Sets and Systems, vol. 327, pp 46-68, 2017.
6. G.Gennady, G.Knyazev: EEG correlates of personality types. Netherlands Journal of Psychology, vol. 62(2), pp 78-87, 2006.
7. M.Grabisch: Set Functions, Games and Capacities in Decision Making. Springer, 2016.
8. A.M.Grass, F.A.Gibbs: A Fourier transform of the electroencephalogram. J. Neurophysiol., vol.1, pp 521–526, 1938.
9. T.Harmony: The functional significance of delta oscillations in cognitive processing. Frontiers in Integrative Neuroscience.<https://doi.org/10.3389/fnint.2013.00083>.
10. D.E.Hinkle, W.Wiersma, S.G.Jurs: Applied Statistics for the Behavioral Sciences. 5th edition. Houghton Mifflin Company, Boston, New York, 2003.
11. G.Knyazev, E.Merkulova, A.Savostyanov, A.Bocharov, A.Saprigyn: Personality and EEG correlates of reactive social behavior. Neuropsychologia, vol. 124, pp 98-107, 2019.
12. E.Pap (editor): Handbook of Measure Theory, two volumes. Elsevier, North Holland, 2002.
13. D.J.Paulus, S.Vanwoerden, P.J.Norton, C.Sharp: From neuroticism to anxiety. Examining unique contributions of three transdiagnostic vulnerability factors. Personality and Individual Differences, vol. 94, pp 38-43, 2016.
14. H.W.Shipton: EEG analysis: a history and a prospectus. Annual Review of Biophysics and Bioengineering, vol. 4, pp 1–15, 1975.
15. B. Stroustrup: The C++ Programming Language. 4th edition. Addison-Wesley Professional, 2013.
16. Z.Wang, G. J. Klir: Generalized Measure Theory. Springer, 2009.
17. Z.Wang, R. Yang, K-S. Leung: Nonlinear Integrals and Their Applications in Data Mining. World Scientific, 2010.