

# Bipolar Rating Scales: A Survey and Novel Correlation Measures Based on Nonlinear Bipolar Scoring Functions

Ildar Batyrshin<sup>1</sup>, Fernando Monroy-Tenorio<sup>1</sup>,  
Alexander Gelbukh<sup>1</sup>, Luis Alfonso Villa-Vargas<sup>1</sup>,  
Valery Solovyev<sup>2</sup>, Nailya Kubysheva<sup>2</sup>

<sup>1</sup> Centro de Investigación en Computación (CIC), Instituto Politécnico Nacional, Av. Juan de Dios Bátiz, C.P 07738, DF, Mexico

<sup>2</sup> Kazan Federal University, 18 Kremlyovskaya Street, Kazan 420008, Russian Federation

batyr1@cic.ipn.mx, b130294@sagitario.cic.ipn.mx, {gelbukh, lvilla}@cic.ipn.mx, {valery.solovyev, NIKubysheva}@kpfu.ru

---

*Abstract: A bipolar rating scale is a linearly ordered set with symmetry between elements considered as negative and positive categories. First, we present a survey of bipolar rating scales used in psychology, sociology, medicine, recommender systems, opinion mining, and sentiment analysis. We discuss different particular cases of bipolar scales and, in particular, typical structures of bipolar scales with verbal labels that can be used for construction of bipolar rating scales. Next, we introduce the concept of bipolar scoring function preserving linear ordering and the symmetry of bipolar scales, study its properties, and propose methods for construction of bipolar scoring functions. We show that Pearson's correlation coefficient often used for analysis of relationship between profiles of ratings in recommender systems can be misleading if the rating scales are bipolar. Basing on the general methods of construction of association measures, we propose new correlation measures on bipolar scales free from the drawbacks of Pearson's correlation coefficient. Our correlation measures can be used in recommender systems, sentiment analysis and opinion mining for analysis of possible relationship between opinions of users and their ratings of items.*

*Keywords: rating scale; bipolar scale; recommender system; opinion mining; sentiment analysis; correlation; association measure*

---

## 1 Introduction

In psychology, sociology, medicine, and other fields, rating scales in numerical or verbal form are often used for measuring human preferences, traits, abilities, and attitudes in order to attribute values of items<sup>1</sup> [37, 20, 22, 23, 19, 12]. Different definitions of rating scale have been proposed. Presently, it is more common to consider rating scale to be a linearly ordered set of categories. The number of categories used in rating scales usually varies from 3 to 11 [20, 22, 17]. A set of rating scales can be aggregated in a scale of higher level [20].

Rating scales often have bipolar structure: two poles and opposite categories symmetrically located at the opposite sides of the scale [37, 20, 23, 27, 24, 3, 35, 11]. In such bipolar scales, the negative side of the scale is the inverse mirror of its positive side [11]. An object is evaluated in bipolar scale as being either positive, negative, or neutral.

In many applications of rating scales, a human is considered as a gauge for measuring and rating the strength of the attitudes or attribute values; the results of such “measurements” are usually represented by numbers. The theory of measurement [25, 30] studies the classes of operations allowed on the sets of such numbers, defining ordinal, interval, and ratio measurement scales. Although rating scales typically have symmetric bipolar form defined by a pair of polar terms [23], this symmetry usually is not considered explicitly in measurement scales. Due to the increasing interest in application of bipolar rating scales in recommender systems and sentiment analysis [28, 1, 29, 21, 34, 9, 26], the problem of consideration of scoring functions defined on bipolar scales explicitly taking into account the symmetry of these scales, which we study in this paper, is of particular interest.

As an alternative to the measurement-based approach to analysis of human attitudes, the model-based approach considers models of words, human verbal evaluations of attributes, sentiment scores, and ratings. For example, in fuzzy logic and soft computing [40, 18], the words representing the strength of the attributes, such as *small* or *very good*, are modeled by fuzzy sets and are often defined by parametric membership functions: triangular, trapezoidal, etc. In fuzzy logic based models of decision making, classification, or control, linguistic terms represented by fuzzy sets are usually considered as components of fuzzy inference systems; these fuzzy sets can be adjusted by a machine-learning procedure for the system to provide the optimal, or reasonable, solutions [18]. Herrera and Herrera-Viedma [16] consider an ordered structure of linguistic terms with the negation operator defined on the set of indexes. The semantics of a linguistic term defined in [16] is given by fuzzy sets defined on the segment  $[0, 1]$ , distributed on it symmetrically or non-symmetrically. Xu [39] considers so-called additive

---

<sup>1</sup> In this paper, we usually list references in chronological order.

linguistic evaluation scales, where scores are given by fractional indexes of linguistic labels and bipolarity of the scale is represented by the negation also defined on the set of indexes. In recommender systems and collaborative filtering, utility function is typically represented by ratings [1]. The rating scales usually have bipolar form, however, generally speaking, utility can be an arbitrary function. Extrapolations from known to unknown ratings in recommender systems are usually done by (a) specifying heuristics that define the utility function and empirically validating their performance, and (b) estimating the utility function that optimizes certain performance criterion, such as the mean square error [1]. Breese et al. [8] consider two alternative probabilistic models for collaborative filtering: Bayesian classifier and Bayesian networks. Machine-learning algorithms are applied to learn parameters of the models. Liu and Seneff [21] consider the problem of assigning scores for the degree of polar sentiment to phrases related with the considered aspects of items. Based on these scores, they calculate an average rating for the aspect. They proposed linear additive model for calculating the scores of polar sentiments taking values in a positive scale. Taboada et al. [34] present lexicon-based approach to extracting sentiment from texts. They use dictionaries of words annotated with their semantic orientation: polarity and strength. The semantic orientation of words takes positive and negative values. It is supposed that some problems of its calculation could be resolved by fine-tuning sentiment orientation values and modifiers [34].

Several online resources have been developed in recent years for sentiment analysis of texts. For example, SenticNet [9] provides polarity scores for 30,000 concepts [26]. The scores range from  $-1$  (bad) to  $+1$  (good), with neutral scores being around zero.

To analyze the similarity between lists of ratings in recommender systems, Pearson's correlation coefficient is often used [29]. However, as we will show, this correlation coefficient can be misleading in analysis of ratings from bipolar scales. Shardanand and Maes [32] proposed to use "constrained" Pearson's coefficient  $r$  to take into account bipolarity of the rating scale.

In this paper, we give a survey of bipolar rating scales. We consider typical structures of bipolar scales with verbal labels that can be used for construction of bipolar rating scales. We introduce the concept of the bipolar scoring function preserving the linear ordering and the symmetry of bipolar scales, study its properties, and propose methods for construction of bipolar scoring functions. We show that Pearson's correlation coefficient often used for analysis of relationships between profiles of ratings in recommender systems can be misleading if the rating scales are bipolar. Basing on general methods of construction of association measures, we propose new correlation measures on bipolar scales free from the drawbacks of Pearson's correlation coefficient.

The paper is organized as follows. In Section 2, we survey different types of bipolar scales discussed in the literature. In Section 3, we present some typical

structures of finite fully-labeled bipolar verbal rating scales. In Sections 4 and 5, we consider formal definitions of bipolar rating scales, bipolar scoring functions, and their properties. In Section 6, we propose novel methods of construction of bipolar scoring functions. In Section 7, we introduce new correlation measures on bipolar rating profiles. Finally, in Section 8 we give some discussion and conclusions.

## 2 Types of Bipolar Scales

Generally speaking, a bipolar scale is a linearly ordered set with minimal and maximal elements considered as two poles of the scale. These poles can be interpreted as negative and positive poles, correspondingly. Usually, bipolar scales have a neutral category located in the center of the scale, and all other categories between the neutral category and the negative and positive poles can be considered as negative and positive categories, correspondingly. In addition to the linear ordering of the bipolar scales, the scale is symmetric: the negative categories are mapped to the corresponding opposite positive categories, and vice versa. The elements of bipolar scales can have numerical scores and linguistic labels.

The symmetry of the bipolar scales can be reflected in the symmetry of the labels or the scores. Typical examples of bipolar scales containing three or five possible responses on questions are considered in [20]. One of these scales with bipolar structure has five categories: *strongly approve*, *approve*, *undecided*, *disapprove*, *strongly disapprove*, with the corresponding scores 5, 4, 3, 2, 1, or scores in reverse order. This scale has a neutral category *undecided* and symmetry between positive and negative categories. Currently, it is more popular to order the scale categories from negative to positive, for example, as follows: *strongly disagree*, *disagree*, *neither agree nor disagree*, *agree*, *strongly agree*, or in correspondence with the ordering of the scores: 1, 2, 3, 4, 5.

Bipolar scales of the following types have been considered in the literature; references given here contain corresponding examples:

- (a) scales with a finite number  $n$  of categories [20],
- (b) scales with infinite number of categories, e.g., in the range from 0 to 1, from  $-1$  to 1, from  $-10$  to 10 [16, 14, 13];
- (c) 2-point scale ( $n = 2$ ) [31] or multipoint scale ( $n > 2$ );
- (d) scales with a neutral category [20] or without neutral category [10];
- (e) scales with verbal labels only at poles [23] or fully labeled when all response categories are explicitly labeled [20, 38];
- (f) scales with symbolic labels of poles [24];

- (g) scales with numeric scores explicitly given [20] or not, e.g., given graphically with intervals [23];
- (h) scales with positive numeric scores, e.g., 1, ..., 5, or from 0 to 1 [20, 16] or with both negative and positive scores, e.g., from -1 to 1 [14, 9];
- (i) scales without polarity of verbal labels; e.g., one of the Likert rating scales [20] contains responses: *grade school, junior high school, high school, college, graduate and professional school* with the scores 1, 2, 3, 4, 5 typical for bipolar scales;
- (j) scales with polarity of verbal categories, but without symmetry with respect to the neutral category, i.e., when some positive verbal category or concept has no corresponding opposite negative verbal category [9, 26];
- (k) scales explicitly using negation operation [40, 16, 39] or not [20];
- (l) scales with non-numeric scores, e.g., with fuzzy sets [40, 16];
- (m) scales with verbal categories ordered from negative categories on the left to positive categories on the right, or in reverse order [20];
- (n) scales with numeric scores increased from left to right: 1, ..., 5 or in reverse order: 5, ..., 1 [20];
- (o) scales with scoring function being linear [20] or nonlinear [34, 39] with respect to the indexes of the categories;
- (p) scales with several verbal labels or concepts assigned to one gradation of the scale, due to synonymy or labels having equal score value [9, 29];
- (q) positive and negative categories separated in two scales [11, 36];
- (r) scales in which an attitude or attribute can have a positive and a negative polarity degree at the same time, e.g., degree of membership and degree of non-membership [2, 11];
- (s) scales with aggregation of scores from the same scale [37, 14, 4];
- (t) scales with aggregation of scores from different scales [20, 21];
- (u) scales with calculation of the group opinion as an aggregation of the scores of individuals [37];
- (v) scales with calculation of individual scores in higher-level scale as aggregation of individual scores in particular scales [20];
- (w) scales measuring [20] or modeling [16, 1] attitudes, preferences, ratings, or utilities.

Below, we consider finite bipolar scales with neutral element as linearly ordered sets with symmetry given by the negation operation. In the following section, we will discuss typical structures of finite bipolar scales fully labeled with verbal labels. Further, we will formally consider two types of mutually related finite bipolar scales with neutral element as sets of integer indexes of categories of bipolar scales. We introduce bipolar scoring functions defined on these sets as models of user preferences.

### 3 Bipolar Verbal Rating Scales

A verbal rating scale is *bipolar* if it is symmetric with respect to the opposite categories, located on the opposite sides of the scale. Such symmetry can be expressed by a *negation operation*  $N$  defined on the scale. Consider an example of a 5-point bipolar scale with verbal labels ordered from left to right: *never* < *seldom* < *sometimes* < *often* < *always* and with negation  $N(\textit{never}) = \textit{always}$ ,  $N(\textit{seldom}) = \textit{often}$ ,  $N(\textit{sometimes}) = \textit{sometimes}$ ,  $N(\textit{often}) = \textit{seldom}$ ,  $N(\textit{always}) = \textit{never}$ . This negation operation has several formal properties. It is *involutive*, i.e., double negation of any category gives the same category, e.g.,  $N(N(\textit{seldom})) = \textit{seldom}$ . It is *decreasing*, i.e., if a category  $c_2$  has greater rating than a category  $c_1$ :  $c_1 < c_2$ , then the negation of  $c_1$  has greater rating than the negation of  $c_2$ :  $N(c_1) > N(c_2)$ . For example, *never* < *often* implies  $N(\textit{never}) = \textit{always} > N(\textit{often}) = \textit{seldom}$ . Note that this negation operation differs from the linguistic *not* [21, 34].

If the scale has an odd number of gradations, then there is a *center, midpoint, or neutral category*  $C$  in the scale such that  $N(C) = C$ . Generally, such point is called the *fixed point* of the negation. For the scale considered above, we have  $C = \textit{sometimes}$ .

Consider a typical structure of 7-point verbal bipolar scale [3]:

$$L = (eap, vap, ap, np, p, vp, ep). \quad (1)$$

$L$  is ordered from left to right:  $eap < vap < \dots < vp < ep$ , where  $p$  and  $ap$  (*anti-p*) denote opposite adjectives or attributes, correspondingly, and other letters denote:  $e = \textit{extra}$ ,  $v = \textit{very}$ ,  $n = \textit{neutral}$ . The negation defines a mapping between the opposite categories, for example:  $N(eap) = ep$ ,  $N(vap) = vp$ , ...,  $N(ep) = eap$ , with  $N(np) = np$  for the neutral category  $np$ . Here is an example of a bipolar verbal scale with the structure (1):

$$L = (\textit{awful}, \textit{very bad}, \textit{bad}, \textit{neither good nor bad}, \textit{good}, \textit{very good}, \textit{excellent}),$$

with  $p = \textit{good}$ ,  $ap = \textit{bad}$ . One can consider the structure (1) as a cupboard with shelves with the labels  $eap, vap, \dots, ep$  ordered from the bottom upward. In a particular application, each “shelf” of (1) can be filled by corresponding verbal labels. These shelves can have scores 1, 2, 3, 4, 5, 6, 7 or  $-3, -2, -1, 0, 1, 2, 3$ , with neutral category  $np$  having the score 4 or 0, respectively. In the following sections, we will consider these scores as indexes of the categories of the bipolar scale with  $n > 3$  indexes, and the score function will be given as a numeric function defined on the set of these indexes. For identity scoring function, its values will coincide with the indexes, in our case, with 1, ..., 7 or with  $-3, \dots, 3$ . Generally, it will be a nonlinear function preserving the symmetry of the bipolar scale.

The shelves of (1) can contain several verbal labels, which are considered synonymous or having equal scores. For example, the shelf  $eap$  can contain the

verbal labels *awful* and *terrible*. A 5-point bipolar scale can be obtained from the structure (1) by eliminating symmetric categories or by symmetric merging of neighboring categories. In such way, one can reduce the structure (1), for example, to the following one:

$$L = (vap, ap, np, p, vp), \quad (2)$$

with indexes 1, 2, 3, 4, 5 or  $-2, -1, 0, 1, 2$ . In the reduced cupboard (2), the shelf *vp* can contain the verbal labels *very good*, *excellent*, *perfect*, etc. Deletion of the neutral category *np* will give bipolar verbal scale without center.

Another typical structure of the bipolar scale is as follows [3]:

$$L = (hap, map, lap, np, lp, mp, hp), \quad (3)$$

where *h*, *m*, and *l* denote *high*, *middle* and *low* intensities, respectively, of the opposite attributes *ap* and *p*. The categories of the scale (3) are ordered from left to right:  $hap < map < \dots < mp < hp$ . The negation defines a mapping between the opposite categories:  $N(hap) = hp$ ,  $N(map) = mp$ , ...,  $N(hp) = hap$ . Bipolar scales with the number of categories less than 7 can be obtained from (3) by deleting pairs of opposite categories.

Generally, the scale structures (1) and (3) can be extended until 15-points bipolar verbal scale by adding gradations with modifiers *el* (*extra low*), *vl* (*very low*), *vh* (*very high*), *eh* (*extra high*) for adjective *p* and its opposite *ap* as follows:

$$L = (ehap, vhap, hap, map, lap, vlap, elap, np, elp, vlp, lp, mp, hp, vhp, ehp). \quad (4)$$

Bipolar scales with the number of categories less than 15 can be obtained from (4) by deleting pairs of opposite categories. Below is an example of such 9-point bipolar scale:

$$L = (\textit{dislike extremely}, \textit{dislike very much}, \textit{dislike moderately}, \textit{dislike slightly}, \textit{neither like nor dislike}, \textit{like slightly}, \textit{like moderately}, \textit{like very much}, \textit{like extremely}),$$

with the center  $C = \textit{neither like nor dislike}$  and with the reduced structure of (4):

$$L = (ehap, vhap, map, lap, np, lp, mp, vhp, ehp).$$

An example of a bipolar verbal 13-point scale can be found in [15].

The considered structures are common for the scales used in sociology, psychology, and medicine. However, in opinion mining and in recommender systems the bipolar scales with symmetry of opposite categories can have categories expressed in a variety of forms. Below is Ringo's scale for rating music [32] in music recommendation system, which is explicitly different from the structures (1) and (3) but implicitly has symmetric form, as in (1), with the center at the category 4:

$L = (1. \text{ Pass the earplugs. } 2. \text{ Barely tolerable. } 3. \text{ Eh. Not really my thing. } 4. \text{ Doesn't turn me on, doesn't bother me. } 5. \text{ Good Stuff. } 6. \text{ Solid. They are up there. } 7. \text{ BOOM! One of my FAVORITE few! Can't live without it}).$

The considered typical structures of bipolar scales can be used for construction of verbal bipolar scales with symmetry of opposite categories. The verbal labels of these categories can differ from one application to another, but the symmetric structure should be preserved in order to consider the verbal rating scales as bipolar scales.

In the next section, we will give a formal definition of a finite bipolar scale as an ordered set of indexes of categories of a bipolar scale considered above. In the sequel, we will consider bipolar scoring functions that assign numerical values to each point of the bipolar scale, which will be used as a numerical model of this scale. The main properties of bipolar utility functions will also be connected with the negation operation related with the symmetry in the scores assigned to opposite categories of the bipolar scale.

## 4 Finite Bipolar Scales

Formally, a *bipolar scale*  $L$  with  $n$  ordered categories  $c_1 < \dots < c_n$  can be represented by an ordered set of indexes of these categories  $J = \{1, \dots, n\}$ ,  $n > 1$ , with the *negation* operation  $N: J \rightarrow J$  defined by

$$N(j) = n + 1 - j \text{ for all } j \in J. \quad (5)$$

The negation function (5) is a *strictly decreasing*:

$$N(i) > N(j) \text{ if } i < j, \quad (6)$$

and *involution*:

$$N(N(j)) = j, \text{ for all } j \in J. \quad (7)$$

We will assume that the bipolar scale has an odd number of elements, i.e.,  $n = 2m + 1$  for some positive integer  $m$ . In this case, the set  $J$  will have a *fixed point* of the negation  $N$ , i.e., an element  $C$  such that

$$N(C) = C. \quad (8)$$

The property (8) is fulfilled for a unique element

$$C = m + 1, \quad (9)$$

called the *center*, *neutral*, or *midpoint* of the bipolar scale. From (5) and (6), it follows that a bipolar scale with an even number of elements  $n$  has no center. Bipolar scales without center can be obtained from scales with center by deleting the center.



For example, the 5-point bipolar scale (*never, seldom, sometimes, often, always*) can be given by an ordered set of indexes  $J = \{1, 2, 3, 4, 5\}$  and with the negation  $N(j) = 6 - j$ , such that  $N(1) = 5, N(2) = 4, N(3) = 3, N(4) = 2, N(5) = 1$ . This scale has a center  $C = 3$ .

From (5),  $n = 2m + 1$ , and (9), we obtain *bipolarity* properties:

$$N(j) + j = 1 + n,$$

and

$$N(j) + j = 2C, \text{ for all } j \in J. \quad (10)$$

Due to bipolarity, the elements  $N(j)$  and  $j$  of the scale are symmetrically located with respect to the center and the “poles” 1 and  $n$  of the scale  $J$ :

$$|j - C| = |N(j) - C|, \quad |j - 1| = |N(j) - n|, \text{ for all } j \in J.$$

We call the ordered set  $K = \{-m, \dots, -1, 0, 1, \dots, m\}$  the *centered form* of the bipolar scale  $J = \{1, \dots, 2m + 1\}$ ,  $m > 0$ . The negation operation  $N: K \rightarrow K$  on  $K$  is defined by:

$$N(k) = -k, \quad \text{for all } k \in K. \quad (11)$$

Unless it can cause confusion, we will use the same letter  $N$  for the negation on  $J$  and on  $K$ , using the arguments  $j$  or  $k$ , respectively. It is clear that  $N$  on  $K$  is a strictly decreasing and involutive function, i.e.,  $N(N(k)) = k$ , for all  $k \in K$ . This scale has the center  $C = 0$  with  $N(C) = C = 0$ , and the bipolarity (10) also fulfills for the scale  $K$  with negation (11):

$$N(k) + k = 2C, \quad \text{for all } k \in K. \quad (12)$$

We call the scale  $K = \{-m, \dots, m\}$ ,  $m > 0$ , with the negation defined by (11) a *centered bipolar scale*.

For example, the 5-point bipolar scale  $J = \{1, 2, 3, 4, 5\}$  is represented in centered form as  $K = \{-2, -1, 0, 1, 2\}$  with the negation  $N$  on  $K$  defined by  $N(-2) = 2, N(-1) = 1, N(0) = 0, N(1) = -1, N(2) = -2$  and with the center  $C = 0$ .

The bipolar scales  $J = \{1, \dots, 2m + 1\}$ ,  $m > 0$ , and  $K = \{-m, \dots, m\}$  can be transformed one into the other as follows:

$$k = j - m - 1, \quad j = k + m + 1, \quad \text{for all } j \in J \text{ and } k \in K. \quad (13)$$

## 5 Bipolar Scoring Functions

Here we study the properties of scoring (utility) functions defined on bipolar scales. The scoring functions give possibility to model users with different utility of categories of the same scale. For example, one user prefers to use the

gradations of the scale near the poles but another user prefers to use the gradations near the center. For such users, the utility of the same gradations can be different. In addition, in the development of decision-making or recommender systems using bipolar scales, the utility of categories of the scales can be represented by non-linear functions modeling utility of categories in different manner, depending on the application or the task. For such functions, the difference between the neighboring gradations depends on their positions on the scale. For example, for the bipolar scale (*never, seldom, sometimes, often, always*), the difference between the utility of the categories *sometimes* and *often* can be modeled by the number 10, but the difference between the utilities of the categories *often* and *always* can be modeled by the number 50. Such nonlinear utility functions can be used, for example, for modeling a user's ratings in model-based approach to collaborative filtering in recommender systems [1].

In model-based approach to construction of bipolar utility functions, how the user measures these utilities is not important; what matters is how to model these utilities for different users and different tasks. In this approach, instead of the "adequate" measurement of a user's preferences or ratings, we concentrate on the effectiveness of the recommendations or decisions generated by the recommender or decision-making system using these bipolar utility functions. Generally, these utility functions can be parameterized and further tuned or adjusted by some machine-learning procedure for the system to obtain optimal or useful recommendations and decisions. For this, these utility functions should satisfy properties similar to the properties of the bipolar scales that we considered above. In this section, we will consider these properties of bipolar utility functions, and in the following section we will consider different methods of construction of such functions.

Let  $I$  denote a bipolar scale  $J$  or  $K$  with  $n = 2m + 1$  categories,  $m > 0$ . For  $J = \{1, \dots, 2m + 1\}$ , we have  $C = m + 1$ ,  $N(j) = n + 1 - j$  for all  $j \in J$ , and for the scale  $K = \{-m, \dots, -1, 0, 1, \dots, m\}$ , we have  $C = 0$ ,  $N(k) = -k$ , for all  $k \in K$ . For both scales, we have  $N(C) = C$ . We call the values  $P_1 = \min(I)$  and  $P_2 = \max(I)$  the negative and the positive poles, correspondingly. For the scale  $J$ , we have  $P_1 = 1$ ,  $P_2 = n$ , and for the scale  $K$ , we have  $P_1 = -m$ ,  $P_2 = m$ . For both scales, we have  $N(P_1) = P_2$  and  $N(P_2) = P_1$ .

**Definition 1.** Let  $I$  be a bipolar scale with negation  $N$ . A strictly increasing real function  $U: I \rightarrow R$  is called a *scoring* or *utility function* on  $I$ . This function is called a *bipolar scoring function* (BSF) on  $I$  if it satisfies the condition

$$U(N(i)) + U(i) = U(P_1) + U(P_2), \quad \text{for all } i \in I. \quad (\text{bipolarity}) \quad (14)$$

For the scales with the center  $C$  from  $N(C) = C$ , we have  $U(P_1) + U(P_2) = 2U(C)$  and the bipolarity property can be given by:

$$U(N(i)) + U(i) = 2U(C), \quad \text{for all } i \in I. \quad (\text{bipolarity}) \quad (15)$$

A BSF is called a *centered bipolar scoring function* (CBSF) if  $U(C) = 0$ .

From (14) and (15), we have:

$$|U(j) - U(C)| = |U(N(j)) - U(C)|, \quad |U(j) - U(P_1)| = |U(N(j)) - U(P_2)|, \quad \text{for all } j \in J,$$

i.e., for a bipolar scoring function  $U$ , the utility values of the opposite categories  $U(j)$  and  $U(N(j))$  are at the equal distances from the utility value of the neutral category  $U(C)$  and at the equal distances from the utility values of the poles.

The definition of the centered bipolar scoring function implies

$$U(N(i)) = -U(i), \quad \text{for all } i \in I. \quad (16)$$

For a CBSF defined on a centered bipolar scale  $K = (-m, \dots, m)$ ,  $m > 0$ , we have:  $U(0) = 0$ ,  $U(k) > 0$  if  $k > 0$ ,  $U(k) < 0$  if  $k < 0$ , and

$$U(-k) = -U(k), \quad \text{for all } k \in K. \quad (17)$$

CBSFs give natural models of utility of categories of bipolar verbal rating scales when these categories have negative and positive sentiments. For example, for the 5-point centered bipolar scale  $K = \{-2, -1, 0, 1, 2\}$ , one can define a CBSF  $U(K) = \{-10, -4, 0, 4, 10\}$  preserving the sign and the symmetry of the bipolar scale  $K$ .

**Proposition 1.** If  $U$  is a BSF on  $I$ , then the function  $W: I \rightarrow R$  defined by

$$W(i) = pU(i) + q \quad \text{for all } i \in I, \quad (18)$$

where  $p, q \in R, p > 0$ , is also a BSF on  $I$ .

**Proof.** It is clear that  $W$  is strictly increasing. Bipolarity of  $W$  follows from (18) and bipolarity (15) of  $U$ :  $W(N(i)) + W(i) = pU(N(i)) + q + pU(i) + q = p(U(N(i)) + U(i)) + 2q = p(2U(C)) + 2q = 2(pU(C) + q) = 2W(C)$ .  $\square$

In (18), the parameters  $p$  and  $q$  define scaling and shifting of the utility function  $U$ , correspondingly.

Proposition 1 implies that from any BSF  $U$  one can obtain a CBSF  $U_C$  as follows:

$$U_C(i) = U(i) - U(C) \quad \text{for all } i \in I. \quad (19)$$

From (5) and (9), it follows that the identity function  $U(j) = j$ , for all  $j \in J$ , is a BSF. We call this function the *standard bipolar scoring function* (SBSF). For example, SBSF on 7-point bipolar scale  $J = \{1, 2, 3, 4, 5, 6, 7\}$  has the values  $U(J) = J = \{1, 2, 3, 4, 5, 6, 7\}$ . From this function, applying the transformation (18) with parameter values  $p = 1$  and  $q = -1$ , one can obtain another popular bipolar scoring function  $U(J) = \{0, 1, 2, 3, 4, 5, 6\}$ .

These examples show that the majority of the popular rating scales, including Likert scales [20] and the scales used in recommender systems [32], use SBSFs. In this paper, we are mainly interested in nonlinear BSFs. For example, on 7-point bipolar scale  $J = \{1, 2, 3, 4, 5, 6, 7\}$  with the center  $C = 4$ , a nonlinear bipolar scoring function can be defined as  $U(J) = \{0, 15, 40, 50, 60, 85, 100\}$  with  $U(C) =$

50. From this BSF, applying the linear transformation (19), one can obtain a CBSF  $U_C(J) = \{-50, -35, -10, 0, 10, 35, 50\}$ .

Note that using transformations (13) of the indexes of the bipolar scales  $J = \{1, \dots, 2m + 1\}$  and  $K = \{-m, \dots, m\}$ , one can transform BSF  $U_J: J \rightarrow R$  defined on  $J$  into a BSF  $U_K: K \rightarrow R$  defined on  $K$ , and vice versa, as follows:

$$U_J(j) = U_K(j - m - 1), \quad U_K(k) = U_J(k + m + 1), \quad \text{for all } j \in J \text{ and } k \in K. \quad (20)$$

These functions have the same set of values:  $U_J(J) = U_K(K)$ . For example, the CBSF  $U_K(K) = \{-10, -4, 0, 4, 10\}$  defined on  $K = \{-2, -1, 0, 1, 2\}$  is transformed into CBSF  $U_J(J) = \{-10, -4, 0, 4, 10\}$  defined on  $J = \{1, 2, 3, 4, 5\}$ . Such simple transformation can be used in the method of construction of CBSF on  $J$  considered in the following section by defining a CBSF as some odd function on  $K$  and then changing the set of its arguments (indexes)  $K$  by  $J$ .

## 6 Methods for Construction of Bipolar Scoring Functions

As we have shown in the previous section, a BSF on a bipolar scale is a linear transformation of an odd function with zero argument value in the center of the scale. The majority of the traditional rating scales can be considered as particular cases of BSFs when the scoring function is linear. In this section, we consider various heuristic methods of construction of bipolar utility (scoring) functions which are, generally, nonlinear.

### 6.1 Bipolar Utility Functions Based on the Distribution of a User's Ratings

We suggest a heuristic method for constructing bipolar utility functions using the distribution of bipolar scores obtained from a user's ratings. Our method is based on the observation that the greater the difference between the frequencies of two neighboring categories, the greater the difference in the utilities of these categories.

For simplicity, consider the categories in a centered bipolar scale  $K = \{-m, \dots, -1, 0, 1, \dots, m\}$ . Suppose the user has provided ratings of  $N$  items with frequencies  $P = (P_{-m}, \dots, P_{-1}, P_0, P_1, \dots, P_m)$ , with  $P_{-m} + \dots + P_m = N$ , where  $P_k$ ,  $k \in \{-m, \dots, m\}$ , is the frequency of the category from the scale  $K$ . Our algorithm consists of the following steps.

1. Symmetrize the frequencies:

$$PS_k = \frac{P_k + P_{-k}}{2}, \text{ for all } k = -m, \dots, m.$$

2. Calculate the cumulative distribution function:

$$U_{-m} = 0, \quad U_{k+1} = U_k + c/|PS_{k+1} - PS_k| + d, \quad k = -m, \dots, m - 1, \tag{21}$$

where  $c$  is a scaling constant and  $d = 0$  if  $PS_{k+1} \neq PS_k$  for all  $k = -m, \dots, m - 1$ , otherwise  $d > 0$ ; the constant  $d$  is introduced for all categories to have different scores and thus the utility function to be strictly increasing.

3. Normalize the utility function to obtain the values of utility function in the range  $[0, M]$ , or  $[-M, M]$ ,  $M > 0$  and, if necessary, move it from the scale  $K = \{-m, \dots, m\}$  to the scale  $J = \{1, \dots, 2m + 1\}$  by replacing indexes by (20).

It can be shown that the proposed method constructs a BSF satisfying the bipolarity property.

Figure 1 shows an artificial example of construction of bipolar utility functions based on the distribution  $P = (5, 35, 48, 60, 30, 20, 2)$  of categories of 7-point bipolar scale in  $N = 200$  ratings of a user. The left plot shows the original and symmetrized distributions and right plot shows the bipolar utility function constructed from the symmetrized distribution. We used  $d = 0$  in (21) because in the symmetrized distribution all neighboring categories have different frequencies.

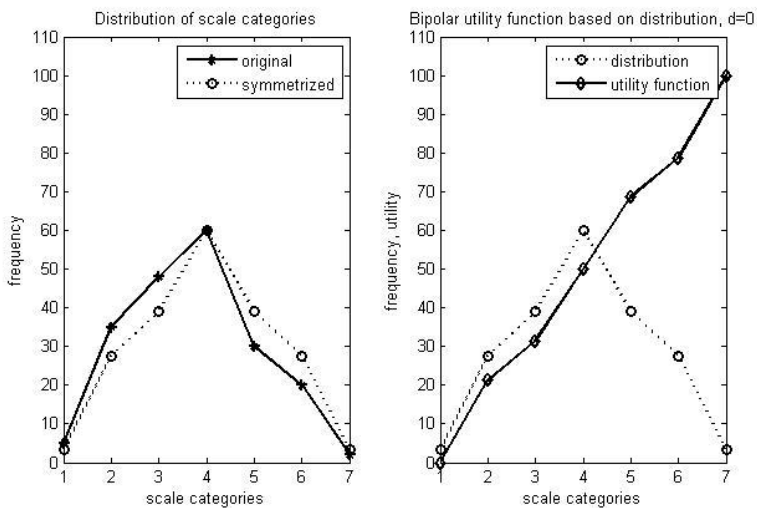


Figure 1.

Construction of bipolar utility function based on the symmetrized distribution of bipolar scale categories obtained in  $N = 200$  ratings of a user, with  $d = 0$

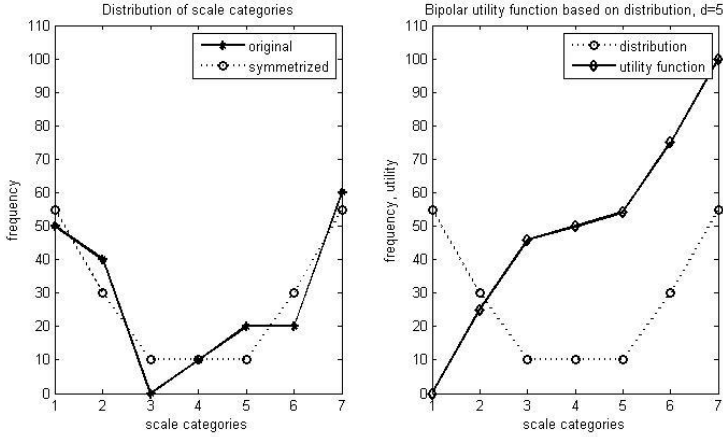


Figure 2

Construction of bipolar utility function based on symmetrized distribution of bipolar scale categories obtained in  $N = 200$  ratings of a user,  $d = 5$

Similarly, in Figure 2 the distribution is  $P = (50, 40, 0, 10, 20, 20, 60)$ . In this case, we used  $d = 5$  because the symmetrized distribution contains neighboring categories with equal frequencies. In both cases, we used  $c = 1$ . All bipolar utility functions were normalized to have values in the interval  $[0, 100]$ .

## 6.2 Generator-Based Utility Functions

Now we suggest a method for constructing CBSF  $U$  on the bipolar scale  $K = \{-m, \dots, -1, 0, 1, \dots, m\}$ ,  $m > 0$ . A bipolar utility function on  $J$  can be constructed using (18) and (20). Let  $G$  be a positive real value and  $g: \{0, \dots, m\} \rightarrow [0, G]$  a strictly increasing function such that  $g(0) = 0$ ,  $g(m) = G$ . We call this function a *generator* of the function  $W: K \rightarrow [-G, G]$  defined by:

$$W(k) = g(k) \quad \text{for all } k \in \{0, \dots, m\}, \quad (22)$$

$$W(k) = -g(-k) \quad \text{for all } k \in \{-m, \dots, -1\}. \quad (23)$$

Obviously, the function  $W$  constructed by this method is a CBSF. In the model-based approach to modeling user preferences, we can define the generator  $g$  by a parametric function and use it to define parametric bipolar utility functions on  $K$  or on  $J$ .

Here are two examples of parametric generators of bipolar utility functions inspired by the parametric Sugeno negation used in fuzzy logic [33], for  $p > -1$ :

$$g_1(k) = \frac{Gk(1+p)}{m+pk}, \quad k = 0, \dots, m, \quad (24)$$

$$g_2(k) = \frac{Gk}{m + pm - pk}, \quad k = 0, \dots, m. \quad (25)$$

Figure 3 shows the shapes of CBSF  $W$  on a 7-point centered bipolar scale  $K = (-3, -2, -1, 0, 1, 2, 3)$  obtained from the generators (24) and (25) for different values of the parameter  $p$ . For  $p = 0$ , both generators are linear:  $g_1(k) = g_2(k) = Gk/m$ .

The generator  $g_1(k)$  with positive  $p$  and the generator  $g_2(k)$  with negative  $p$  can be used for modeling bipolar utility functions when the scores are located nearer to the boundaries of the scale; see Figure 3. For such utility functions, the nearer the categories of the scale to the corresponding positive or negative pole, the smaller the difference between the utilities of the neighboring categories. For example, if on the 7-point scale

$L = (\text{awful}, \text{very bad}, \text{bad}, \text{not good not bad}, \text{good}, \text{very good}, \text{excellent})$

we define a bipolar utility function with these properties, then the difference between the utilities of the categories *very good* and *excellent* will be small.

Conversely, the generator  $g_1(k)$  with negative  $p$  and the generator  $g_2(k)$  with positive  $p$  can be used for modeling bipolar utility functions when the scores are located near the center of the scale, i.e., the nearer the categories of the scale to the corresponding pole, the larger the difference between the utility values of the neighboring categories. For example, if on the same scale  $L$  we define a function with these properties, then the utility value of the category *excellent* will be much greater than that of the category *very good*.

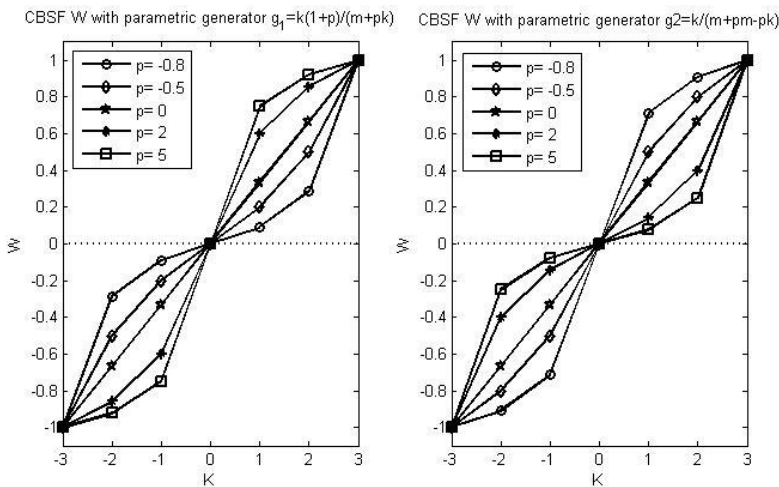


Figure 3

Examples of CBSF on 7-point centered bipolar scale generated by generators  $g_1(k)$  (left) and  $g_2(k)$  (right) for different values of parameter  $p$

In the model-based approach to analysis of user opinions, the parameters of the utility functions can be tuned for the model to generate the best solutions.

### 6.3 Selection of Generator Values

A general method of construction of the generator of bipolar utility function given by (22) and (23) can be summarized as follows:

```

g(0)=0; g(m) = G;
for k = 1 to m - 1:
    Determine u ∈ (g(k - 1), G);
    g(k) = u;
end

```

The function *Determine* depends on the specific method, for example: random selection, selection based on optimization of some criteria, etc.

## 7 Correlation Measures on the Set of Bipolar Profiles

Pearson's product-moment correlation coefficient

$$\text{corr}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (26)$$

is often used in recommender systems for measuring similarity between profiles [29]. In this section, we show that this correlation coefficient can be misleading if the opinions are measured in bipolar scales. Consider the following utility profiles with the ratings of 10 items in 7-point bipolar scale  $J = \{1, 2, 3, 4, 5, 6, 7\}$  with the standard utility function  $U(j) = j$  for all  $j$  in  $J$ :

$$x = (7, 5, 5, 7, 7, 7, 5, 7, 5, 5),$$

$$y = (5, 7, 7, 5, 5, 5, 7, 5, 7, 7),$$

$$z = (3, 1, 1, 3, 3, 3, 1, 3, 1, 1).$$

The profiles  $x$  and  $y$  have only “positive” (greater than neutral  $C = 4$ ) ratings, so a reasonable association measure  $A$  should show positive association between them:  $A(x, y) > 0$ ; however, the correlation coefficient gives  $\text{corr}(x, y) = -1$ . The profiles  $x$  and  $z$  have almost opposite (“positive” vs. “negative”) ratings, so a reasonable association measure should give negative association between them:  $A(x, z) < 0$ ; however, the correlation coefficient gives  $\text{corr}(x, z) = 1$ .



Therefore, we need to introduce correlation (association) measures that, similarly to the correlation coefficient, could show positive and negative associations between profiles of ratings in bipolar scales, but without the drawbacks of Pearson's correlation coefficient, such as shown in the above example. Below we present such measures, based on general results discussed in [5–7].

Let  $I$  be a bipolar scale ( $I = J$  or  $I = K$ ) with the negation  $N$  and with the center  $C$ . We call the vector  $x = (x_1, \dots, x_M)$ ,  $x_s \in I$ ,  $s = 1, \dots, M$ , of elements from the bipolar scale  $I$  a *rating profile*. We also call the vector  $C_X = (C, \dots, C)$  of the length  $M$  the *central profile* of the set  $X$  of all rating profiles of the length  $M$ . We define the *negation of the profile*  $x$  as  $N_X(x) = (N(x_1), \dots, N(x_M))$ . Obviously,  $N_X$  is an involution on  $X$ , i.e.,  $N_X(N_X(x)) = x$  for all profiles from  $X$ , and  $C_X$  is a unique fixed point of  $N$ :  $N_X(C_X) = C_X$ . For a bipolar utility function  $U$  defined on the bipolar scale  $I$ , we call the vector  $U_X(x) = (U(x_1), \dots, U(x_M))$  a *utility profile* of the rating profile  $x$ .

Suppose a user evaluates 6 items in the bipolar rating scale  $J = \{1, 2, 3, 4, 5\}$  by the vector of ratings  $x = (3, 5, 2, 4, 1, 3)$ . Suppose  $U(J) = \{-10, -3, 0, 3, 10\}$  is the centered bipolar utility function defined on  $J$ . Then, the utility profile of the rating profile  $x$  is given by  $U(x) = (0, 10, -3, 3, -10, 0)$ , the negation of the profile  $x$  is given by  $N_X(x) = (3, 1, 4, 2, 5, 3)$ , and on the set  $X$  of all rating profiles of the length 6, the central profile is given by  $C_X = (3, 3, 3, 3, 3, 3)$ .

Consider two rating profiles  $x$  and  $y$  with the same length. We will define the correlation measure  $A_U(x, y)$  as a function of utility profiles  $U(x)$  and  $U(y)$ . In applications, when the users profiles have different lengths, the vectors  $x$  and  $y$  will contain only ratings of items presented in the profiles of both users.

**Definition 2.** Let  $X$  be the set of all profiles of the length  $M$  with ratings from the bipolar scale  $I$  with the negation  $N$  and the center  $C$ . Let  $U$  be a bipolar utility function defined on  $I$ . A *correlation (association) measure* on the set  $V = X \setminus \{C_X\}$  is a function  $A_U: V \times V \rightarrow [-1, 1]$  that satisfies for all  $x, y \in V$  the following properties:

$$A_U(x, y) = A_U(y, x) \quad (\text{symmetry}) \quad (27)$$

$$A_U(x, x) = 1, \quad (\text{reflexivity}) \quad (28)$$

$$A_U(x, N(y)) = -A_U(x, y). \quad (\text{inverse relationship}) \quad (29)$$

We call a correlation measure  $A_U$  *C-separable* if it satisfies the following properties:

$$A_U(x, y) > 0 \text{ if for all } s = 1, \dots, M \text{ it holds } x_s, y_s > C \text{ or } x_s, y_s < C, \quad (30)$$

$$A_U(x, y) < 0 \text{ if for all } s = 1, \dots, M \text{ it holds } y_s < C < x_s \text{ or } x_s < C < y_s. \quad (31)$$

The properties (27)–(29) were used in [7] in the definition of the correlation (association) measures on the set  $X$  with involution  $N$ . These properties generalize the properties of Pearson's correlation coefficient applied to  $M$ -tuples when the

negation of  $M$ -tuples of real values is defined by  $N(x) = -x = (-x_1, \dots, -x_M)$ . Here we extend the definition of association measures given in [7] on the set of bipolar utility profiles. The properties (30) and (31) are introduced here to avoid the problems with the correlation coefficient defined on bipolar profiles discussed at the beginning of this section. See also  $C$ -separability property of association measures on  $[0,1]$  considered in [6].

From (28) and (29), we have

$$A_U(x, N(x)) = -1. \quad (32)$$

Definition 2 can be extended from the set  $V$  to the set of all profiles  $X$  replacing the property (28) by

$$A_U(x, x) = 1 \text{ if } x \neq C_X \quad (33)$$

In this case, for all  $x$  in  $X$  we have

$$A_U(x, C_X) = A_U(C_X, x) = 0. \quad (34)$$

Consider a method for construction of correlation (association) measures on the set of bipolar utility profiles based on the general methods discussed in [5].

**Proposition 2.** Let  $I$  be a bipolar scale ( $I = J$  or  $I = K$ ) with the center  $C$ ,  $X$  be a set of profiles  $x = (x_1, \dots, x_M)$  of the length  $M$ ,  $x_s \in I$ ,  $s = 1, \dots, M$ , with the central profile  $C_X = (C, \dots, C)$ , and  $U$  be a bipolar utility function on  $I$ . Then the following function is a  $C$ -separable correlation measure on  $X \setminus \{C_X\}$ :

$$A_U(x, y) = \frac{1}{2^t} \sum_{s=1}^M [ |F(x_s) + F(y_s)|^t - |F(x_s) - F(y_s)|^t ], \quad (35)$$

where

$$F(x_s) = \frac{U(x_s) - U(C)}{\sqrt[t]{\sum_{s=1}^M |U(x_s) - U(C)|^t}}, \quad (36)$$

and  $t \geq 1$ .

One can easily check that the properties (27)–(31) are satisfied for the function (35).

If the bipolar scoring function  $U$  in (36) is centered, i.e.,  $U(C) = 0$ , then (36) is simplified as follows:

$$F(x_s) = \frac{U(x_s)}{\sqrt[t]{\sum_{s=1}^M |U(x_s)|^t}}. \quad (37)$$

Proposition 2 implies the following corollary.

**Corollary 1.** In the conditions of Proposition 2, if  $t = 2$  in (35), (36), then the following function is a  $C$ -separable correlation measure on  $X \setminus \{C_x\}$ :

$$A_U(x, y) = \frac{\sum_{s=1}^M (U(x_s) - U(C))(U(y_s) - U(C))}{\sqrt{\sum_{s=1}^M (U(x_s) - U(C))^2} \sqrt{\sum_{s=1}^M (U(y_s) - U(C))^2}}. \quad (38)$$

If the bipolar scoring function  $U$  in (38) is centered, then:

$$A_U(x, y) = \cos(U(x), U(y)) = \frac{\sum_{s=1}^M U(x_s)U(y_s)}{\sqrt{\sum_{s=1}^M U(x_s)^2} \sqrt{\sum_{s=1}^M U(y_s)^2}}, \quad (39)$$

where  $U(x) = (U(x_1), \dots, U(x_M))$ ,  $U(y) = (U(y_1), \dots, U(y_M))$ .

Since the formulas (37) and (39) require fewer operations than (36) and (38) in calculation of the correlation value  $A_U(x, y)$  between a large number of pairs  $(x, y)$  of profiles, it is recommended to replace a bipolar utility function  $U$  defined on the bipolar scale  $I$  by a centered bipolar utility function  $U - U(C)$  and then calculate the correlation between the corresponding profiles by (35), (37) or by (39).

Let us calculate correlation  $A_U(x, y)$  between the profiles of ratings  $x = (7, 5, 5, 7, 7, 7, 5, 7, 5, 5)$ ,  $y = (5, 7, 7, 5, 5, 5, 7, 5, 7, 7)$ ,  $z = (3, 1, 1, 3, 3, 3, 1, 3, 1, 1)$  from the bipolar scale  $J = \{1, 2, 3, 4, 5, 6, 7\}$  considered at the beginning of this section. For the standard utility function  $U(J) = J$ , we obtain  $U(x) = x$ ,  $U(y) = y$ ,  $U(z) = z$ . We can calculate correlation between three profiles by (38). However, as we have noted above, it is more efficient to replace the bipolar utility function  $U$  by the centered utility function  $U_C(J) = U(J) - U(C) = J - 4 = K = \{-3, -2, -1, 0, 1, 2, 3\}$  and to use this centered utility function in (39). We obtain:  $A_U(x, y) = 0.6$ ,  $A_U(y, z) = -1$ ,  $A_U(x, z) = -0.6$ . These values correspond to our propositions  $A(x, y) > 0$  and  $A(x, z) < 0$  considered at the beginning of this section, and the new correlation measure (38) does not have the drawbacks of Pearson's correlation coefficient. Note that we have  $z = N(y)$ , and for this reason, according to the property (29) of the correlation measure, we obtain:  $A_U(x, z) = A_U(x, N(y)) = -A_U(x, y) = -0.6$ ; according to the property (32), we obtain  $A_U(y, z) = A_U(y, N(y)) = -1$ .

The correlation measure (38) obtained here as a particular case of (35) generalizes the constrained correlation coefficient considered in [32] (see formula (5)) using in (38) the standard 7-point utility function  $U = J$  with the center  $C = 4$  and  $U(C) = 4$ . Consider (38) in the form (39), when the utility function  $U$  is replaced by its centered form  $U - U(C)$ . As one can see, the formula (39) is a nonlinear function sensitive to the presence of respectively high utility values for the same items in both profiles  $x$  and  $y$ .

Consider an example for 7-point scale:  $K = \{-3, -2, -1, 0, 1, 2, 3\}$ . Suppose one uses the standard utility function  $U(K) = K$ , and two users have the following profiles of rating of four items in the scale  $K$ :  $x = \{1, 1, 1, 1\}$ ,  $y = \{-1, -1, -1, -1\}$ . Due to  $U(K) = K$ , the corresponding utility profiles will have the same values  $U(x) = x$ ,  $U(y) = y$ . Since the profiles are opposite:  $N(x) = y$ , the correlation (39) between them has the value  $A(x, y) = -1$ . Suppose a new item has the rating 2 from both users. Then, we obtain the new profiles:  $x^* = \{1, 1, 1, 1, 2\}$ ,  $y^* = \{-1, -1, -1, -1, 2\}$  with the correlation value  $A(x^*, y^*) = 0$ . As one can see, addition of the same value 2 to both ratings drastically changes the correlation value from  $-1$  to 0. This situation can be avoided if we use nonlinear bipolar utility function. Let  $U(1) = 1$  and  $U(2) = 1.5$ . Then, we obtain the following utility profiles:  $U(x^*) = \{1, 1, 1, 1, 1.5\}$ ,  $U(y^*) = \{-1, -1, -1, -1, 1.5\}$  with correlation between them  $A_U(x^*, y^*) = -0.28$ . As one can see, the change of the correlation value from  $-1$  to  $-0.28$  is not as drastic for considered nonlinear bipolar utility function as for the standard linear utility function.

Consider another example of profiles with ratings from 7-point rating scale  $K$ . Suppose again that we use the standard utility function  $U(K) = K$  and two users have equal rating profiles of five items:  $x = \{1, 1, 1, 1, 2\}$ ,  $y = \{1, 1, 1, 1, 2\}$ . We have  $A(x, y) = 1$ . Suppose for the sixth item both users have the opposite ratings 3 and  $-3$ . For the standard utility function, the correlation between new utility profiles  $x^* = \{1, 1, 1, 1, 2, 3\}$  and  $y^* = \{1, 1, 1, 1, 2, -3\}$  is drastically changed, from the value  $A(x, y) = 1$  to the value  $A(x^*, y^*) = -0.059$ . Let us change the standard utility function by  $U(K) = \{-2, -1.5, -1, 0, 1, 1.5, 2\}$ . For this nonlinear bipolar utility function, the correlation between utility profiles changes from  $A(x, y) = 1$  to  $A_U(x^*, y^*) = 0.22$ , which is not as drastic as for the standard utility function.

In the two considered examples, we used the method of construction of bipolar utility functions considered in Section 6.3, which defines the positive part of the centered utility function and symmetrically maps it to the negative part of the scale with 0 in the center of the scale. Similarly, we defined sequentially  $U(1) = 1$ ,  $U(2) = 1.5$ , and  $U(3) = 2$ . Another method can be based on a parametric generator such as (24) or (25). For example, using the generator (24) with the parameter  $p = 1$ , we can obtain the similar results: for nonlinear bipolar utility function  $U(K) = \{-2, -1.6, -1, 0, 1, 1.6, 2\}$ , the correlation between utility profiles  $x^*$  and  $y^*$  has the value  $A_U(x^*, y^*) = 0.24$ .

In both examples, we decreased the absolute utility values of the categories near the poles to avoid the drastic change of the correlation value when both users use near polar ratings for the same items. In general, a parametric utility function can be heuristically adjusted to obtain intuitively validated solutions or optimized by some machine-learning method for the recommender or decision-making system using bipolar rating scales to obtain solutions with better performance.

## 8 Discussion and Conclusions

The rating scales applied in different application areas usually have bipolar structure, but this bipolarity between opposite categories of the scales often was not explicitly or formally exploited. Our paper introduces explicitly the property of bipolarity in the definition of the general structure of verbal bipolar scales, in the formal definition of the bipolar scale as the linearly ordered set of indexes with negation operation and in the definition of bipolar scoring function on bipolar scale preserving the symmetry of this scale. In the description of the general structure of bipolar scales in Section 3 we based on the paper [3]. The idea of the formal definition of bipolar scale on the set of indexes was partially based on the paper [16] where the linguistic categories of the scale are presented by fuzzy sets. In Section 4 we consider together two mutually related sets of indexes where  $J = \{1, \dots, 2m+1\}$  is more traditional and  $K = \{-m, \dots, m\}$  is more “natural” for representation of bipolar scales with opposite categories. Most of bipolar scales observed in Sections 1 and 2 can be represented as bipolar scales with general structure considered in Section 3 or formally as bipolar scales considered in Section 4. The concept of bipolar scoring or utility function defined on bipolar scale to the best of our knowledge is new. The bipolar scoring functions include the traditional scoring of  $n$ -point rating scales by numbers  $1, \dots, n$  as particular case and such scoring functions are called standard bipolar scoring functions. But it seems more interesting instead of the standard scoring functions or instead of the linear utility functions to consider nonlinear utility functions. These utility functions can be given as parametric functions and adjusted by some machine learning procedure to obtain good or optimal results on the output of recommender or decision making system using these bipolar scales. The nonlinear bipolar scoring functions can be useful in modeling the ratings of users or in modeling utility or importance of categories in bipolar scales. The reasons to use nonlinear bipolar utility functions in correlation measure introduced in the paper are discussed in Section 7. The results on association measures considered in Section 7 are based on the papers [5-7]. Here, we use the terms association measure and correlation measure as interchangeable. We extend the property of  $C$ -separability from association measure on  $[0,1]$  considered in [6] on the set of utility profiles. The general formula for association measure on bipolar utility profiles is based on Minkowski distance and on general results considered in [5] for time series. The formulas (37) and (39) for centered bipolar utility functions are specific for  $C$ -separable association measures on bipolar utility profiles. The formula (38) generalizes the constrained correlation coefficient considered in [32] without utility functions. In our future work we plan to apply the results of the paper in collaborative filtering and in analysis of human ratings.

### Acknowledgements

This work was supported in parts by the projects 20162204 and 20171344 of SIP IPN, 240844 and 283778 of CONACYT, 15-01-06456A of RFBR and by the

Russian Government Program of Competitive Growth of Kazan Federal University.

### References

- [1] Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), pp. 734-749.
- [2] Atanassov, K. T. (1986). Intuitionistic fuzzy sets. *Fuzzy sets and Systems*, 20 (1), 87-96.
- [3] Batyrshin, I.Z. (1990). On the structure of verbal scales. In: *Proceedings of the Second All-Union Conference on Artificial intelligence*. Minsk, 1990, vol. 1, pp. 37-40.
- [4] Batyrshin, I. (2011). Uncertainties with memory in construction of strict monotonic t-norms and t-conorms for finite ordinal scales: basic definitions and applications. *Applied and Computational Mathematics*, vol. 10, 3, 2011, pp. 498-513.
- [5] Batyrshin, I. (2013). Constructing time series shape association measures: Minkowski distance and data standardization. In *2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence (BRICS-CCI & CBIC)*, (pp. 204-212). IEEE. <https://arxiv.org/abs/1311.1958v3>.
- [6] Batyrshin, I. Z. (2015). Association measures on  $[0, 1]$ . *Journal of Intelligent & Fuzzy Systems*, 29(3), pp. 1011-1020.
- [7] Batyrshin, I. Z. (2015). On definition and construction of association measures. *Journal of Intelligent & Fuzzy Systems*, 29(6), 2015, pp. 2319-2326.
- [8] Breese, J. S., Heckerman, D., & Kadie, C. (1998, July). Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence* (pp. 43-52). Morgan Kaufmann Publishers Inc.
- [9] Cambria, E., Olsher, D., & Rajagopal, D. (2014). SenticNet 3: a common and common-sense knowledge base for cognition-driven sentiment analysis. In *Proceedings of the twenty-eighth AAAI conference on artificial intelligence* (pp. 1515-1521). AAAI Press.
- [10] Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, 18(3), 205-215.
- [11] Dubois, D., & Prade, H. (2006). Bipolar representations in reasoning, knowledge extraction and decision processes. In *International Conference*

- on Rough Sets and Current Trends in Computing, pp. 15-26. Springer Berlin Heidelberg.
- [12] Friberg, O., Martinussen, M., & Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences*, 40(5), 873-884.
- [13] Goldberg, K., Roeder, T., Gupta, D., & Perkins, C. (2001). Eigentaste: A constant time collaborative filtering algorithm. *information retrieval*, 4(2), 133-151.
- [14] Grabisch, M. (2006). Aggregation on bipolar scales. In *Theory and applications of relational structures as knowledge instruments II* (pp. 355-371). Springer Berlin Heidelberg.
- [15] Gunderman, R. B., & Chan, S. (2013). The 13-point Likert scale: a breakthrough in educational assessment. *Academic radiology*, 20(11), 2013, pp. 1466-1467.
- [16] Herrera, F., & Herrera-Viedma, E. (2000). Linguistic decision analysis: steps for solving decision problems under linguistic information. *Fuzzy Sets and systems*, 115(1), 67-82.
- [17] Hjerstad, M. J., Fayers, P. M., Haugen, et al (2011). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *Journal of Pain and Symptom Management*, 41(6), 1073-1093.
- [18] Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice Hall.
- [19] Juniper, E. F., Guyatt, G. H., Ferrie, P. J., & King, D. R. (1999). Development and validation of a questionnaire to measure asthma control. *European Respiratory Journal*, 14(4), 902-907.
- [20] Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*. Vol 22, No. 140, 1932, pp. 55. New York.
- [21] Liu, J., & Seneff, S. (2009). Review sentiment scoring via a parse-and-paraphrase paradigm. In *Proceed. 2009 Conference on Empirical Methods in Natural Language Processing: Vol. 1* (pp. 161-169). Association for Computational Linguistics.
- [22] Mosier, C.I. (1941). A psychometric study of meaning. *The Journal of Social Psychology*, 13(1), 123-140.
- [23] Osgood, C.E. (1952). The nature and measurement of meaning. *Psychological bulletin*, 49(3), 197-237.

- [24] Petrenko V.F. (1988). *Psychosemantics of consciousness*. Moscow, Ripol Klassik. (In Russian: Петренко, В. Ф. (1988). Психосемантика сознания. Рипол Классик).
- [25] Pfanzagl, J. (1971). *Theory of measurement*. Physica. Physica-Verlag Heidelberg.
- [26] Poria S, Gelbukh A, Cambria E, Hussain A & Huang G (2014). EmoSenticSpace: A novel framework for affective commonsense reasoning, *Knowledge-Based Systems*, 69, pp. 108-123.
- [27] Pospelov D.A. (1989). *Models of Reasoning. Essay in the Analysis of Mental Acts*. Radio y Svyaz. Moscow. (In Russian: Пospelов Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов, М.: Радио и связь, 1989, -184с.)
- [28] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp. 175-186. ACM.
- [29] Ricci, F., Rokach, L., Shapira, B., Kantor P.B. (2011) *Recommender Systems Handbook*, Springer US.
- [30] Roberts, F. S. (1985). *Measurement theory*. Cambridge University Press.
- [31] Schafer, J. H. J. B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. *The adaptive web*, 291-324.
- [32] Shardanand, U., & Maes, P. (1995). Social information filtering: algorithms for automating “word of mouth”. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 210-217). ACM Press/Addison-Wesley Publishing Co.
- [33] Sugeno M. (1974) *Theory of Fuzzy Integrals and its Applications*. Dissertation. Tokio Institute of Technology.
- [34] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307.
- [35] Tarassov V.B. (2001). Analysis and modeling of NOT-factors on polar scales. In: *Integrated Models and Soft Computing in Artificial Intelligence*, Moscow, Fismatlit, pp. 65-71. (In Russian: Тарасов В.Б. Анализ и моделирование НЕ-факторов на полярных шкалах// Интегрированные модели и мягкие вычисления в искусственном интеллекте. – М.: Наука. Физматлит, 2001. – С.65-71).
- [36] Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12), 2544-2558.



- [37] Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology*, 33(4), 529-554.
- [38] Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27(3), 236-247.
- [39] Xu, Z. (2012). *Linguistic decision making*. Springer Berlin Heidelberg.
- [40] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning.—I. *Information sciences*, 8(3), 1975, pp. 199-249; —II. *Information sciences*, 8(4), 1975, pp. 301-357. —III. *Information sciences*, 9(1), 1975, pp. 43-80.