Decision Making in Social Media with Consistent Data

José I. Peláez[a,*], Eustaquio A. Martínez[b], Luís G. Vargas[c]

[a]*Department of Languages and Computer Sciences. IBIMA University of Málaga, 29071 Málaga, Spain*
[b]*Polytechnic Faculty, East National University, Ciudad del Este, 7000, Alto Paraná, Paraguay*
[c]*Joseph M. Katz Graduate School of Business, University of Pittsburgh, 15260, Pennsylvania, USA*

Abstract

The use of data obtained from social media in decision-making is growing, because people are increasingly using these means to inform themselves, express opinions or make valuations of brands, services, etc. Increasingly, people and organizations use this information to make strategic decisions for their business. This decision process involves, among other thing, obtaining a ranking of alternatives to support the decision. However, in many decision-making models the use of data is done without considering important aspects such as consistency or contextualization, which causes the results to be questioned in a general way due to lack of rigor in the decision-making. In this work, a decision-making model is proposed to obtain an alternative ranking contextualizing the feelings/opinions of the users represented through intervals of feeling consistent, using data obtained from the social media. This model uses an interval majority aggregation operator, which constructs opinion intervals using data extracted from the social media; a consistency index for pairwise-comparison interval matrices; an algorithm to reconstruct consistent interval pairwise-comparison matrices, by means of a deflationary process of intervals' diameter; and an operator to obtain a ranking of alternatives in the interval pairwise-comparison matrices. The model has been applied to a real case showing adequate results according to the market.

Keywords: Reciprocal matrices, Consistency Index, Preference relations, Interval data, Majority Operators, Decision Making

1. Introduction

In recent years, we have observed the development of social networks, which has drastically transformed the way people communicate and obtain information [1–3]. Currently, social networks are present in daily life, playing an increasingly key role in the organizations' management. Progressively, organizations use information from social networks to offer services, interact with their audiences, in short, to make strategic decisions for their businesses. The contents generated by the public, opinions, experiences, feelings, offers opportunities and challenges to organizations. For example, in the commercial field, consumers increasingly consult the opinions generated by other users, to evaluate the products or services before making a purchase. To increase competitive advantage and effectively evaluate the business, organizations need to monitor and analyze the opinions generated by their audiences about their businesses in social

* Corresponding author.
Email addresses: jipelaez@uma.es (José Ignacio Peláez), amartinez@fpune.edu.py (Eustaquio Alcides Martínez), lgvargas@pitt.edu (Luís G. Vargas)

media. Different studies show a marked growth in the performance of organizations that have a strong capacity for business analysis of their audience's opinions [4,5].

Organizations often perform marketing intelligence to collect and analyze data from internal and external information sources. The mechanism for collecting and analyzing these sources is called business intelligence, the part that focuses on internal business data; and competitive intelligence the portion that focuses on external business data [6]. In recent years, due to advances in social media technology, the amount of data from online social networks has grown explosively [1–3]. Leskovec [7] indicates that the content generated by the public in the form of blog posts, comments and tweets establish a connection between organizations and consumers. Therefore, it is expected that organizations take advantage of this data generated by users to extract entities and topics, understand the consumer, visualize relations and create their marketing intelligence to excel in a business environment. In particular, a marketing intelligence report may include market information about the popularity of competitors' products and services, consumer sentiments about their products and services, promotional information and/or activities offered by competitors [8].

One of the challenges for improving business competitiveness is to correctly understand the meaning of the sentiments in the unsolicited opinions of the public in social networks [3,9], to solve problems such as ranking of alternatives in decision-making process. This improvement depends on different factors, among which stand out: (1) the accuracy of the data through more effective semantic analysis processes, to more accurately determine the sentiment of comments made by the public about products /services/...; (2) the representation of the extracted data, to more accurately model the feelings associated with the data [10]; and (3) the consistency of the comments/unsolicited data, so that the decisions are made based on data that have been issued in a conscious manner and not in an ignorant way [11].

Although the improvement in the semantic analysis of data has experienced important advances, applying multi-agent systems incorporating in parallel systems based on neural networks, machine learning, and so on [12,13], the same cannot be said in relation to the representation or consistency of the public unsolicited opinions in social media [10,14]. Most systems extract a feeling value from the comments and then use average values of all comments, usually an arithmetic mean, to obtain the overall value of feeling, and create a ranking of alternatives.

The problem of ranking alternatives has been addressed numerous times in scientific decision, mainly by the implementation of multiple-criteria decision making analysis (MCDA) methods [15–20]. But the way in which all these methods use the information, causes a significant loss of information for decisions [2], because by using a single value instead of using the interval [11,21], that encompasses all the values of feeling, some variation in preferences might somehow be lost as well as the cardinality of preferences that are similar, that somehow reinforce those preferences. Likewise, most methods do not consider that the data has not been issued with ignorance, which would cause decisions to be made with inconsistent data. Only the Analytic Hierarchy process [22], adds an inconsistency index to consider the ignorance in the decision process, but this index is designed to reciprocal matrices.

The objective of this work is development a decision-making model to obtain an alternative ranking, which contextualizes unsolicited opinions expressed by decision-makers in social media, in consistent preference intervals. This model uses an interval majority aggregation operator, which constructs opinion intervals using data extracted from the social media; a consistency index for pairwise-comparison interval matrices; an algorithm to reconstruct consistent interval pairwise-comparison matrices, by means of a deflationary process of intervals diameter; and an operator to obtain a ranking of alternatives in the interval pairwise-comparison matrices.

The paper has been organized as follows: In section 2, we present the background, where some related works are briefly commented and the aggregation operators and inconsistency index, to understand the new decision model. The section 3 presents the decision model, and the new interval consistency index ICI^+ and the aggregation operator AC-OWG, that yields the ranking of the alternatives from an interval pairwise-comparison matrix. In section 4, an example where international telecommunications companies are analyzed, and finally in section 5, the conclusions and future direction of this research.

2. Background

In this section we present briefly some related works, and three relevant issues for the design of the work proposal: the consistency index CI^+, the aggregation operators SMA-OWA, and C-OWG.

2.1. Related works

The problem of ranking alternatives has been addressed several times in decision science, generally through the implementation of multiple-criteria decision making analysis (MCDA) methods [15–20]. The uncertainty experienced by the decision makers when making comparisons was measured, associating each trial with a range of numerical values and analyzing their effects on the change of rank, to obtain the final ranking of the preferences of the alternatives [11], that encompasses all the values of feeling, some variation in preferences might somehow be lost as well as the cardinality of preferences that are similar, that somehow reinforce those preferences. Furthermore, several methods do not consider that the data has not been issued with ignorance, which would cause decisions to be made with inconsistent data. As mentioned, only the Analytic Hierarchy Process [22] adds an inconsistency index to consider the ignorance in the decision process, but this index is designed to reciprocal matrices.

Different modeling approaches to the emotional profile in a specific location to extract and summarize the reviews (opinions) about a product, of all the clients was explored in [3]. Extraction and summarization of the reviews (opinions) about a product, of all the clients was proposed in [23] and the modeling of social networks is carried out by Leskovec in [7].

The main problem of all these models is that they do not consider the variety of opinions regarding a product or service, because they only take one value, which can be the average, to represent the opinion/sentiment, instead of using an interval that represents the variation of opinions or feelings.

2.2. The index CI+

Peláez et al [24] defined the consistency index CI^+ for reciprocal pairwise comparison matrices as follows:

$$CI^+ = \begin{cases} 0, & \text{if } n < 3 \\ CI^+(M_{n\times n}), & \text{if } n = 3 \\ \dfrac{1}{\Omega(M_{n\times n})} \displaystyle\sum_{i=1}^{\Omega(M_{n\times n})} CI^+(\Upsilon_i), & \text{if } n > 3 \end{cases} \quad (1)$$

where:

- $M_{n\times n}$ is the pairwise-comparison matrix,
- Υ_i is the *ith* transitivity of matrix M,
- $\Omega(M_{n\times n}) = (n-3)!\,3!/n!$, is the number transitivity cycles of a matrix.

This index has the advantage that it can be applied to different scales and consistency relations, included the additive (additive in [24]) relationship. CI^{+} determines the consistency of the matrices through the minimum element of consistency, the transitivity cycles.

2.3. Aggregation Operator SMA-OWA

Peláez et al [25–27] defined the aggregation operator SMA-OWA that incorporates the cardinality factor c_i of the elements $v_i > 0$ being aggregated and S_n (aggregate) permutation group and $\sigma \in S_n$, any permutation. It is a mapping $F_{SMA}: \mathbb{R}^n \times \mathbb{N}^n \rightarrow \mathbb{R}$ given by:

$$F_{SMA}(r) = \sum_{i=1}^{n} w_{i,N} v_{\sigma(i)} \quad (2)$$

where: $N = max_{1 \le i \le n} c_i$; $\sigma \in S_n$ is an ordering (aggregate) permutation; the weights $v_{\sigma(i)}$ satisfy the order relation $v_{\sigma(i)} \ge v_{\sigma(i+1)}$, and the weights $w_{i,N}$ are defined by the following recurrence relations:

$$w_{i,1} = u_1 = 1/n \quad (3)$$

$$w_{i,k} = \frac{\gamma_{i,k} + w_{i,k-1}}{u_k}, 2 \le k \le N, 1 \le i \le n \quad (4)$$

$$u_k = 1 + \sum_{j=1}^{n} \gamma_{j,k}$$

and (5)

$$\gamma_{i,k} = \begin{cases} \xi, & c_{\sigma(j) \ge k} \\ 1 - \xi, & Otherwise \end{cases}$$

where ξ is the cardinality relevance factor (CRF), $0 \le \xi \le 1$.

If $\xi = 1$, the value of the majority's opinions is taken into account, if $\xi = 0.5$, the arithmetic average of the opinions is obtained, and if $\xi = 0$, only the minority's opinion is considered [27]

The SMA-OWA aggregation operator models the majority opinion through the cardinality of its elements, through the CRF ξ.

2.4. The C-OWG Operator

Yager and Xu [28,29] defined the C-OWG (Continuous Ordered Weighted Geometric) operator to obtain the ranking of alternatives from an interval reciprocal pairwise-comparison matrix (IPCM) $M_{n \times n} = \{m_{ij} = [m^L{}_{ij}, m^U{}_{ij}]\}$. It is defined as a mapping h_Q from the space of closed intervals with positive lower bounds $\mathrm{I}\mathbb{R}^{+}$ (Real intervals) to $\mathbb{R}^{+}$, with an associated differentiable BUM (Basic Unit-Interval Monotonic [30]) function $Q: [0,1] \rightarrow [0,1]$ with the following properties: $Q(0) = 0$; $Q(1) = 1$ and $Q(x) \ge Q(y)$ if $y > x$:

$$h_Q([m_{ij}^L, m_{ij}^U]) = m_{ij}^U . \left(\frac{m_{ij}^L}{m_{ij}^U}\right)^{\int_0^1 (dQ(y)/dy) y dy} \quad (6)$$

$h_Q(m_{ij})$ yields the expected preference degree value of alternative a_i over the alternative a_j. The overall expected preference degree value of alternative a_i over all the alternatives is given by the geometric mean:

$$g_Q(a_i) = \left(\prod_{j=1}^{n} h_Q(m_{ij})\right)^{1/n}, \quad i = 1,2,\dots,n \quad (7)$$

The greatest value of $g_Q(a_i)$ implies that a_i is the best alternative, and the alternatives a_i $(i = 1,\dots,n)$ can be ranked based on the values $g_Q(a_i)$ $(i = 1,\dots,n)$ [28].

3. Proposed Model

In this section, we present a decision model to rank the alternatives based on consistency users' feelings/opinions about a product or service in the social media. The model will be presented using the Open Group Architecture Framework (TOGAF) [31], that provides a global view of the model from a business perspective, showing the global objectives and the system motivation, the data sources, interfaces and the entire process.

Also, in this section, we present the new aggregation operator ISMA-OWA to construct the interval pairwise-comparison matrix; the interval consistency index ICI^{+}; and the AC-OWC operator to calculate the ranking of alternatives from interval pairwise-comparison matrices.

3.1. The ISMA-OWA Operator

ISMA-OWA is an interval operator from majority-operator family ([25,32]). This operator constructs from a set of discrete values and their cardinalities, a range that represents that set, i.e., given a set of values $r = (r_1, \dots, r_n) \in \mathbb{R}^n \times \mathbb{N}^n$, a permutation group S_n and $\sigma_n \in S_n$ such that $r_{\sigma(1)} \leq \cdots \leq r_{\sigma(n)}$, a function G is searched such that:

$$G(r) = [r^L, r^U],\ r^L, r^U \in \mathbb{R},\ r^L \leq r^U \qquad (8)$$

So that the function is $G: \mathbb{R}^n \times \mathbb{N}^n \to I\mathbb{R}$ (real intervals [21]), $r_i = (v_i, c_i)$ and that:

$$\begin{aligned} r^L &= F_{SMA_\sim}(r_\sim) \\ r^U &= F_{SMA^\sim}(r^\sim) \end{aligned} \qquad (9)$$

where $(r_\sim, r^\sim) = g(r)$, such that:

$$g: \mathbb{R}^n \times \mathbb{N}^n \to \mathbb{R}^\alpha \times \mathbb{N}^\alpha \qquad (10)$$

$$\alpha = \begin{cases} n - \varepsilon, & \text{if } g(r) = r_\sim \\ \varepsilon, & \text{if } g(r) = r^\sim \end{cases}$$

ε depends on the function $g(r)$. We have

$$r^\sim = (r_1, \dots, r_\varepsilon), r_\sim = (r_{\varepsilon+1}, \dots, r_n) = (r_{\sim 1}, \dots, r_{\sim n-\varepsilon}) \qquad (11)$$

$$F_{SMA_\sim}(r_\sim) = \sum_{i=1}^{n-\varepsilon} \omega_{\sim 1,N_\sim}\, v_{\sigma(i)\sim} \qquad (12)$$

$$F_{SMA^\sim}(r^\sim) = \sum_{i=1}^{\varepsilon} \omega^\sim{}_{1,N^\sim}\, v_{\sigma(i)}{}^\sim$$

where:

$N_\sim = max_{1 \leq i \leq n-\varepsilon} c_i$

$N^\sim = max_{1 \leq i \leq \varepsilon} c_i$

$\omega_\sim$ and $\omega^\sim$ are calculated according to (2)-(5) using $v_{\sigma(i)\sim}$, $v_{\sigma(i)}{}^\sim$.

For this operator, it is important to define the CRF ξ, one for each data subset, to obtain the values r^L and r^U of the interval $[r^L, r^U]$ with the representative values of both data sets.

Definition 1: Operator ISMA-OWA is a mapping $F_{ISMA}: \mathbb{R}^n \times \mathbb{N}^n \to I\mathbb{R}$, ($I\mathbb{R}$ is the set of real intervals) defined as follows:

$$\begin{aligned} F_{ISMA}(r) &= \left[F_{SMA_\sim}(r_\sim), F_{SMA^\sim}(r^\sim)\right] \\ &= \left[\sum_{i=1}^{n-\rho} \omega_{\sim 1,N_\sim}\, v_{\sigma(i)\sim}, \sum_{i=1}^{\rho} \omega^\sim{}_{1,N^\sim}\, v_{\sigma(i)}{}^\sim\right] \end{aligned} \qquad (13)$$

where:

$(r_\sim, r^\sim) = g(r)$, such that $g: \mathbb{R}^n \times \mathbb{N}^n \to \mathbb{R}^\alpha \times \mathbb{N}^\alpha$ $(\alpha = n - \rho)$ if $g(r) = r_\sim$; $\alpha = \rho$, if $g(r) = r^\sim$ defined as:

$$g(r) = (\phi_1(r), \phi_2(r)) \qquad (14)$$

where:

$$r^\sim = \phi_1(r) = \begin{cases} r_i, i < \rho \\ \mu(r_i), i = \rho \end{cases}; i = 1, \dots, \rho$$

$$r_\sim = \phi_2(r) = \begin{cases} r_i, i > \rho + 1 \\ r_{\rho+1}, i = \rho + 1 \end{cases}; i = \rho + 1, \dots, n$$

ρ is the index of r cutoff-point.

therefore:

$$\begin{aligned} r^\sim &= (r_1, \dots, \mu(r_\rho)) \\ r_\sim &= (r_{\rho+1}, \dots, r_{n-\rho}) = (r_{\sim 1}, \dots, r_{\sim n-\rho}) \end{aligned}$$

$\mu(r_\rho) = (\hat{v}, \hat{c})$ is the term that represents a statistical measure such as the median or the mean $\hat{v}$ and its cardinality $\hat{c}$.

As an application example of the previously presented operator, suppose we have the feelings set $r = \{0.8, 0.8, 0.6, 0.5, 0.3, 0.3\}$ with its associated cardinality set $C = \{2, 1, 1, 2\}$. Applying the operator ISMA-OWA with CRF $\xi = 1$ (to determine the value of the majority's opinion as upper and lower bounds of the intervals); we obtain the interval [0.35, 0.75]. This interval is obtained as follows: first, we calculate ρ, as the position of the median in the r set ($\rho = 3$). Then, we get the sub-sets $r_{\sim} = \{0.5, 0.3, 0.3\}$ and $r^{\sim} = \{0.8, 0.8, 0.6\}$, and applied the operator $F_{ISMA}(r)$ as $[F_{SMA_{\sim}}(r_{\sim}), F_{SMA^{\sim}}(r^{\sim})]$ to obtain the interval [0.35, 0.75]. This interval represents the variation of feelings regarding a comparison, considering the majority expression for each of the two sets generated from S. In this case, the data median has been used as the cutoff point.

3.2. The Index ICI^+

Based on the index CI^+ [24], here we propose the Interval Consistency Index (ICI^+).

In this work we represent the unsolicited data from social media through intervals, using the operator ISMA-OWA to represent in a range all the feelings/opinions about a product or service. The intervals are used to construct a matrix with additive scale, with entries $m_{ij} \in [0,1]$. The consistency relation to this kind of matrices is given by $m_{ik} = m_{ij} - 0.5 + m_{jk}$, $m_{ii} = 0.5$ and $m_{ji} = [1,1] - [m_{ij}^L, m_{ij}^U]$, then an IPCM is given by:

$$M = [m_{ij}]_{n\times n} = \begin{bmatrix} [0.5,0.5] & [m_{12}^L, m_{12}^U] & \dots & [m_{1n}^L, m_{1n}^U] \\ [1-m_{12}^U, 1-m_{12}^L] & [0.5,0.5] & \dots & [m_{2n}^L, m_{2n}^U] \\ \vdots & \vdots & \ddots & \vdots \\ [1-m_{1n}^U, 1-m_{1n}^L] & [1-m_{2n}^U, 1-m_{2n}^L] & \dots & [0.5,0.5] \end{bmatrix} \quad (15)$$

where do we get M^U and M^L:

$$M^U = \begin{bmatrix} 0.5 & m_{12}^U & \dots & m_{1n}^U \\ 1-m_{12}^L & 0.5 & \dots & m_{2n}^U \\ \vdots & \vdots & \ddots & \vdots \\ 1-m_{1n}^L & 1-m_{2n}^L & \dots & 0.5 \end{bmatrix} \quad (16)$$
$$M^L = \begin{bmatrix} 0.5 & m_{12}^L & \dots & m_{1n}^L \\ 1-m_{12}^U & 0.5 & \dots & m_{2n}^L \\ \vdots & \vdots & \ddots & \vdots \\ 1-m_{1n}^U & 1-m_{2n}^U & \dots & 0.5 \end{bmatrix}$$

By separating the consistency relation from its interval format, we have $m_{ik}^L = m_{ij}^L - 0.5 + m_{jk}^L$ and $m_{ik}^U = m_{ij}^U - 0.5 + m_{jk}^U$, that corresponds to the consistency relations of M^U and M^L, respectively. Then, according to [33], M would have an acceptable consistency if M^U and M^L were simultaneously consistent. Consequently, considering the consistency index CI^+, it is easy to see that CI^+ can be extended to define an interval consistency index ICI^+ applicable to an IPCM, according to the following definition.

Definition 2. The interval consistency index ICI^+ of an interval positive reciprocal matrix $M_{n\times n}$ is defined as:

$$ICI^+(M) = \begin{cases} 0 & ;\text{if } n < 3 \\ \min\{CI^+(M^L), CI^+(M^U)\} & ;\text{if } n = 3 \\ \min\left\{ \frac{1}{\Omega(M_{n\times n}^L)} \sum_{i=1}^{\Omega(M_{n\times n}^L)} CI^+(\Upsilon_i^L), \frac{1}{\Omega(M_{n\times n}^U)} \sum_{i=1}^{\Omega(M_{n\times n}^U)} CI^+(\Upsilon_i^U) \right\} & ;\text{if } n > 3 \end{cases} \quad (17)$$

When $n > 3$, ICI^+ it is calculated as the average of the transitivity cycles existing in order 3 matrices, which can be obtained from M^L and M^U, when their minors are calculated through the main diagonal until all the order 3 matrices are found [34]. Then, since $0 \le CI^+(M) \le 1$, matrix M is consistent if and only if $\delta_* \le ICI^+(M) \le 1$, where δ_* is the critical acceptance value that is determined through percentiles [24]. This critical value can be easily extended for an IPCM, so that $\delta_* \le ICI^+(M)$ ensures the or acceptable consistency of M^L and M^U simultaneously, since $min\{\delta_1, \delta_2\}$ must be greater than or equal to δ_*.

3.3. The AC-OWG Operator

AC-OWG (Additive continuous ordered weighted geometric) is an additive interval aggregation operator to calculate the ranking of alternatives from interval pairwise-comparison matrices. This operator is an adaptation of the C-OWG [28], to handling data with an additive preference relation.

In an interval additive preference relation, an IPCM is consistent if $[m_{ik}^L, m_{ik}^U] = [m_{ij}^L, m_{ij}^U] - [0.5,0.5] + [m_{jk}^L, m_{jk}^U]$ and $[m_{ji}^L, m_{ji}^U]$ =[1,1]- $[m_{ij}^L, m_{ij}^U]$, for all i and j. These relations lead to $m_{ik}^L = m_{ij}^L - 0.5 + m_{jk}^L$, $m_{ik}^U = m_{ij}^U - 0.5 + m_{jk}^U$, $m_{ji}^L = 1 - m_{ij}^U$, and $m_{ji}^U = 1 - m_{ij}^L$. Then, according to (6), we have:

$$h_Q(m_{ij}) = m_{ij}^U \left(\frac{m_{ij}^L}{m_{ij}^U}\right)^{\int_0^1 (dQ(y)/dy)ydy} =$$

$$m_{ij}^{U\,1-\int_0^1 (d(Q(y)/dy)ydy} m_{ij}^{L\,\int_0^1 (d(Q(y)/dy)ydy}.$$

from [35] $\int_0^1 Q(y)dy = 1 - \int_0^1 (dQ(y)/dy)ydy$ and hence, then we have:

$$h_Q(m_{ij}) = m_{ij}^{U\,\int_0^1 Q(y)dy} m_{ij}^{L\,\int_0^1 (dQ(y)/dy)ydy} =$$

$$= m_{ij}^{U\,\int_0^1 Q(y)dy} m_{ij}^{L\,1-\int_0^1 Q(y)dy} = m_{ij}^L \left(\frac{m_{ij}^U}{m_{ij}^L}\right)^{\int_0^1 Q(y)dy}$$

$$\text{i.e., } h_Q(m_{ij}) = m_{ij}^L \left(\frac{m_{ij}^U}{m_{ij}^L}\right)^{\int_0^1 Q(y)dy} \quad (18)$$

Since $m_{ji} = [1,1] - m_{ij} = [1 - m_{ij}^U, 1 - m_{ij}^L]$ then we have:

$$h_Q(m_{ji}) = h_Q([1 - m_{ij}^U, 1 - m_{ij}^L]) =$$

$$(1 - m_{ij}^U)\left(\frac{1-m_{ij}^L}{1-m_{ij}^U}\right)^{\int_0^1 Q(y)dy} =$$

$$h_Q(m_{ji}) = m_{ji}^L \left(\frac{m_{ji}^U}{m_{ji}^L}\right)^{\int_0^1 Q(y)dy} \quad (19)$$

Then, the AC-OWG is defined according to the following:

Definition 3. The Operator AC-OWG is a mapping $\psi: \mathbb{IR}^+ \to \mathbb{R}^+$given by:

$$AC - OWG(m_i) = \left(\prod_{j=1}^{n} \psi_Q(m_{ij})\right)^{1/n}, \quad (20)$$
$$i = 1,2, \dots, n$$

with an associated BUM (basic unit-interval monotonic) function $Q: [0,1] \to [0,1]$ with the following properties: $Q(0) = 0$; $Q(1) = 1$ and $Q(\text{x}) \geq Q(\text{y})$ if $y > x$, such that:

$$\psi_Q(m_{ij}) = \psi_Q([m_{ij}^L, m_{ij}^U]) = m_{ij}^L \left(\frac{m_{ij}^U}{m_{ij}^L}\right)^{\lambda},$$

$$\psi_Q(m_{ji}) = \psi_Q([m_{ji}^L, m_{ji}^U]) = m_{ji}^L \left(\frac{m_{ji}^U}{m_{ji}^L}\right)^{\lambda}, \quad (21)$$

$$\forall\, i \leq j$$

where λ is the attitude character of Q, given by:

$$\lambda = \int_0^1 Q(y)dy.$$

3.4. The Proposed Model

Figure 1 shows the business model described in this paper. It is composed of three large blocks: objectives and motivation, data sources and interfaces, and the process. consistent data, with the main restriction of using unsolicited information from social media. The objective of the Business Model is to build a ranking of the alternatives from Data Sources and Interfaces, the sole purpose of this component is to provide data to the Process block.

The stakeholders, e.g., companies, marketing agencies, etc., oversee identifying the alternatives they wish to analyze to determine consumers' purchasing decisions. By means of their social media communications, consumers provide information to determine pairwise comparisons.

by audiences in social media. Intervals using ISMA-OWA operator constructs the feeling interval for each pair of alternatives, thus building the IPCM. Consistency analysis using ICI^+ determines whether the matrix is consistent. If it is consistent, it is passed to the next step. However, if it is inconsistent, a deflationary process is performed in the intervals, to search for a consistent matrix, in case it exists; otherwise the critical value of accepted matrices can be modified (acceptance percentile is softened) or new data sets are requested, and the process starts again. Finally, the aggregation using the AC-

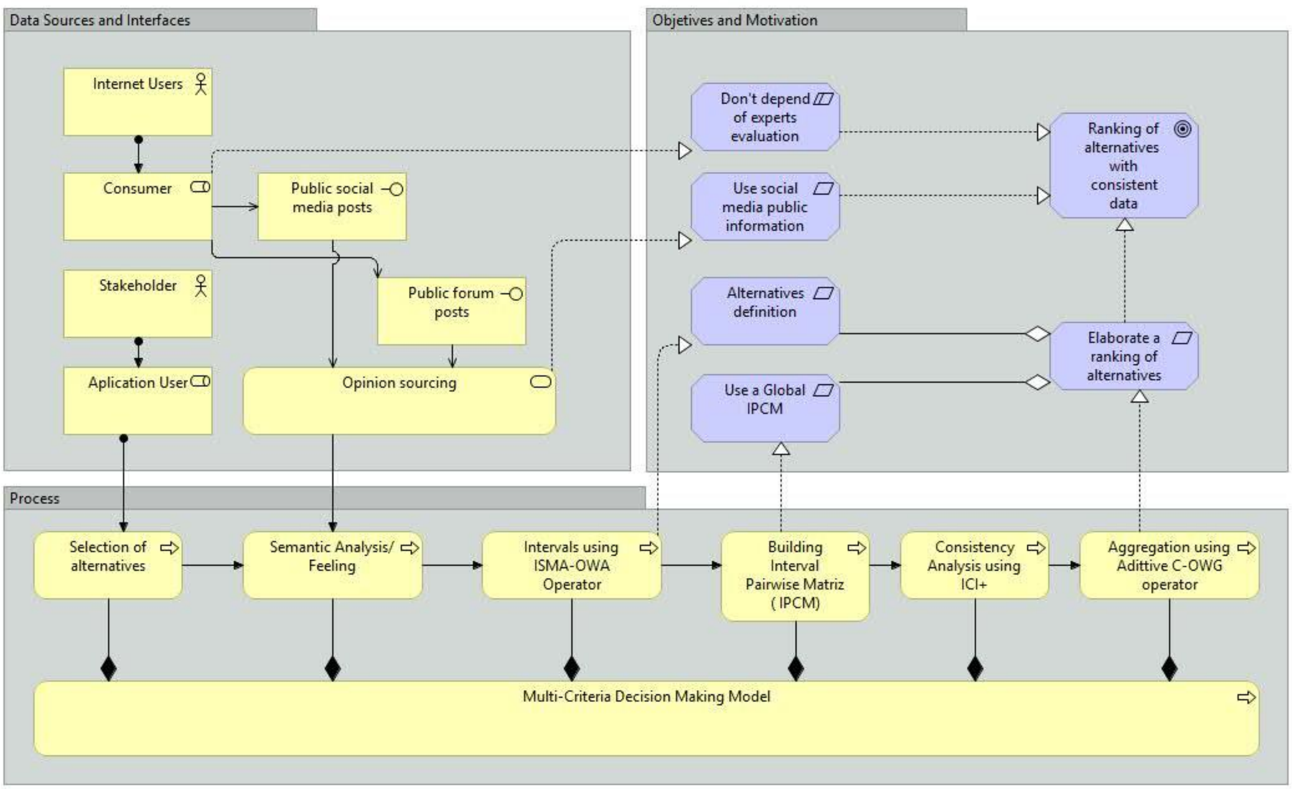


Figure 1 Proposed Decision-making model

The process block is made up of a decision-making model that yields the ranking of the alternatives. The flow of the model is as follows: the process begins with the Selection of Alternatives, where the client specifies which products he/she wants to compare. Semantic Analysis/Feeling calculates the comparisons polarity (positive, neutral, negative, 1, 0.5 and 0 respectively) expressed

OWG operator is applied to obtain the ranking of the alternatives.

3.4.1. *Extraction and Analysis of Digital Ecosystem Data*

Information can be extracted from social media through several alternatives that may include algorithms that use typical social media APIs such as the one provided by Twitter, News Services, or Web Scraping processes (Web Crawling, Screen Scraping, Web Extraction, Crawl Spider, Web-Bot, Spider Robot, Data Mining, Harvests, etc.) [23,36,37]. These communications analysis requires algorithms of semantic analysis, which allows for the extraction of users' feelings and at the same time determines each of the communication topics related to the alternatives [23,36–39].

Selecting the sources, compiling the information and its localization should guarantee the validity of the information [40]. To design the model, a process has been implemented to locate information sources by means of web crawlers [41].

For this work, a system has been implemented that identifies relevant communications and analyzes them. The process can be observed in the architecture shown in figure 2.

stakeout, that communication is stored, to be later analyzed using the Python NLTK [39] library and a set of logical rules that define the actor to be analyzed, that is, to analyze the technical service of a company, for example: *"The tech support guy of company **A** was very nice*" any occurrence type *"company **A**"* or *" Company **A**"* and *"Tech support"*, *"Company **A**"*, *"Tech support"*, etc. They must be considered. Since only communications that match the contextualization pattern are analyzed and can be modified dynamically.

However, when the communications obtained are not relevant enough, precision can be achieved by imposing stricter restrictions. Also, when changing the contextualization rules, the same analysis can be done in different alternatives or in a subset of its communications, that is, "*Company **A***" and "*Tech support*"; that allows valuations in each one of the defined criteria.

To represent the sentiment polarity of the communications in the interval [0,1], a multi-

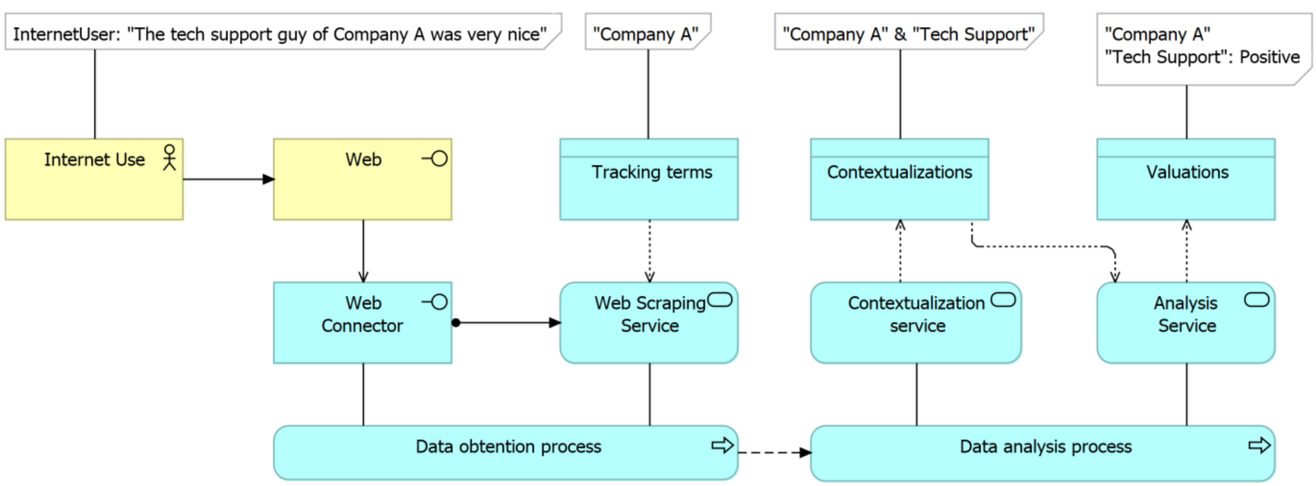


Figure 2. Extracting alternatives from Social Media architecture

A step prior to the analysis is fundamental and consists of the definition of terms that are sought or followed. These terms are usually words, a set of words or sentences that a system monitoring service will verify in its communication search task. Then, when some user in a social network under verification through its API integrated into the system is manifested, using some of the terms under

agent system has been implemented that incorporates two algorithms. The first one based on neural networks, and the second one based on machine learning [12,13,39].

3.4.2. Interval Pairwise Comparison Matrix Building

The IPCM, with additive scale, is constructed using the opinions and/or feeling related to a product or service obtained from social networks, according the process explained at section 3.4.1. These opinions or feelings are previously processed to obtain a value between 0 and 1, and on these values and their cardinalities the ISMA-OWA operator is applied, which allows obtaining an interval representing the set of opinions regarding to the product or service evaluated, comparing two alternatives a_i and a_j. Then, the comparison is represented by the respective interval $m_{ij} = [m_{ij}^L, m_{ij}^U]$. Here an important point arises, although the values obtained from social networks are discrete, the intervals obtained by the ISMA-OWA operator contain the variety of those discrete values, which is the why this transformation is assumed to be valid. An example of IPCM with the mentioned characteristics and constructed as explained is shown in table 1.

Table 1: Interval pairwise-comparison matrix with additive scale.

$Alt.$	a_1	a_2	⋯	a_n
a_1	[0.5,0.5]	$[m_{12}^L, m_{12}^U]$	⋯	$[m_{1n}^L, m_{1n}^U]$
a_2	$[1-m_{12}^U, 1-m_{12}^L]$	[0.5,0.5]	⋯	$[m_{2n}^L, m_{2n}^U]$
⋯	⋮	⋮	⋱	⋮
a_n	$[1-m_{1n}^U, 1-m_{1n}^L]$	$[1-m_{2n}^U, 1-m_{2n}^L]$	⋯	[0.5,0.5]

$Alt.$ = Alternatives

3.4.3. Making the IPCM Consistent

A IPCM M, can be consistent or inconsistent. If M is inconsistent (M^L or M^U or both inconsistent), the intervals of its entries are deflated using $[m_{ij}^L + \Delta, m_{ij}^U - \Delta]$, until a consistent IPCM is found (see algorithm 1). That is to enclosure in the new IPCM a crisp acceptable consistent matrix [42]. If a consistent matrix cannot be found, and the intervals have reached the minimum acceptable diameter (i.e., the lower and upper bounds are the same or the lower bound is larger than upper bound), we can proceed in two different ways: (1) modifying the critical acceptance value, thus relaxing the acceptance percentile [24]; or (2) introducing new data and calculating new intervals with more information. This process is shown in Algorithm 1.

Algorithm 1. Process for consistency analysis of the IPCM

Algorithm: Deflation_Matrices_IPMC
Input : Public's Preferences
Output: M matrix type IPCM
Step 1: Construct M using ISMA-OWA ($\xi = 1$)
Step 2: Verify M consistency using ICI^+
Step 3: If M is consistent Then go to step 5
Else deflate the intervals of M, $[m_{ij}^L + \Delta, m_{ij}^U - \Delta]$, and return to step 2. If the IPCM entries can no longer be deflated, because they have reached the data limits (you can no longer, get new interval) go to step 4.
Step 4: Change the acceptance percentile, go to step 5, or get another data set and go to step 1.
Step 5: Return (M).

3.4.4. Determining the Ranking of Alternatives from Interval Matrices.

When a consistent IPCM is obtained employing algorithm 1, to determine then rank of alternatives, the Operator AC-OWG is used, according to (20).

4. Application's Example

The purpose of this example is to illustrate how the proposed model works. We have considered a real example with companies and social media communications related to the telecommunications sector in Spain. The objective is to obtain an emotional ranking of the main international companies operating in this sector: MoviStar, Orange, Jazztel and Vodafone.

Sample Information:
Companies:
- MoviStar,
- Orange,
- Jazztel,
- Vodafone.

Considered data sources:
Social media in general
Period: January-August, year 2016.

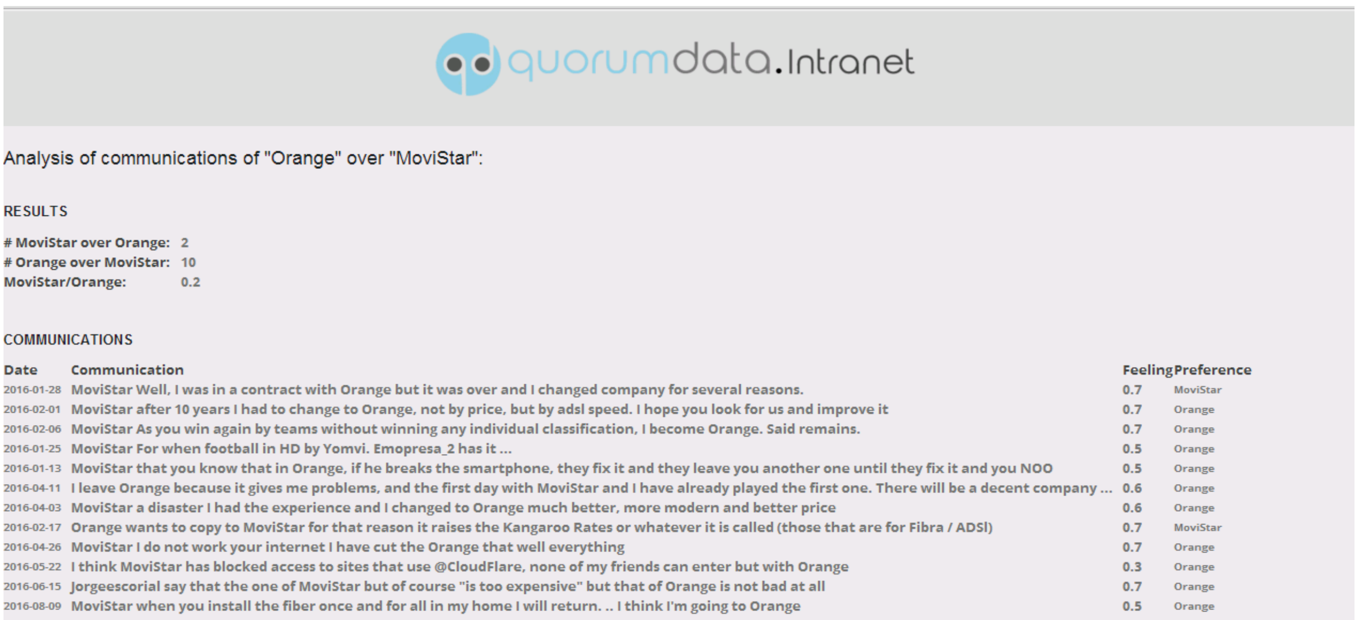
Total Comments: 60,000 pairwise-comparison comments, for each two alternatives.

- Movistar-Orange : 10,000
- Movistar-Jazztel : 10,000
- Movistar-Vodafone : 10,000
- Orange-Jazztel : 10,000
- Orange-Vodafone : 10,000
- Jazztel-Vodafone : 10,000

Feeling scale: [0,1].

Step 1. Information Extraction and Feeling Analysis

In Figure 3, we give an excerpt of the comments used for the study along with their sentiment values. Once the feeling of each communication is obtained, they are grouped by these values and their cardinality, obtaining tuples that contain values and cardinality of each feeling.



quorumdata.Intranet

Analysis of communications of "Orange" over "MoviStar":

RESULTS

 # MoviStar over Orange: 2
 # Orange over MoviStar: 10
MoviStar/Orange: 0.2

COMMUNICATIONS

Date	Communication	Feeling	Preference
2016-01-28	MoviStar Well, I was in a contract with Orange but it was over and I changed company for several reasons.	0.7	MoviStar
2016-02-01	MoviStar after 10 years I had to change to Orange, not by price, but by adsl speed. I hope you look for us and improve it	0.7	Orange
2016-02-06	MoviStar As you win again by teams without winning any individual classification, I become Orange. Said remains.	0.7	Orange
2016-01-25	MoviStar For when football in HD by Yomvi. Emopresa_2 has it ...	0.5	Orange
2016-01-13	MoviStar that you know that in Orange, if he breaks the smartphone, they fix it and they leave you another one until they fix it and you NOO	0.5	Orange
2016-04-11	I leave Orange because it gives me problems, and the first day with MoviStar and I have already played the first one. There will be a decent company ...	0.6	Orange
2016-04-03	MoviStar a disaster I had the experience and I changed to Orange much better, more modern and better price	0.6	Orange
2016-02-17	Orange wants to copy to MoviStar for that reason it raises the Kangaroo Rates or whatever it is called (those that are for Fibra / ADSl)	0.7	MoviStar
2016-04-26	MoviStar I do not work your internet I have cut the Orange that well everything	0.7	Orange
2016-05-22	I think MoviStar has blocked access to sites that use @CloudFlare, none of my friends can enter but with Orange	0.3	Orange
2016-06-15	Jorgeescorial say that the one of MoviStar but of course "is too expensive" but that of Orange is not bad at all	0.7	Orange
2016-08-09	MoviStar when you install the fiber once and for all in my home I will return. .. I think I'm going to Orange	0.5	Orange

Figure 3. Example of some communications used for the study of two companies..

Step 2. Construction of Pairwise-Comparison Matrix

The pairwise comparison of additive interval values is constructed by applying the ISMA-OWA operator to the feeling tuples. The interval matrix for the example is shown in Table 2.

Table 2: interval pairwise-comparison matrix examples for telecommunications sector.

M

$$= \begin{pmatrix} \textbf{Company} & MoviStar & Orange & Jazztel & Vodafone \\ MoviStar & [0.500, 0.500] & [0.300, 0.500] & [0.300, 0.500] & [0.499, 0.699] \\ Orange & [0.499, 0.700] & [0.500, 0.500] & [0.498, 0.700] & [0.499, 0.698] \\ Jazztel & [0.499, 0.699] & [0.299, 0.501] & [0.500, 0.500] & [0.499, 0.700] \\ Vodafone & [0.300, 0.500] & [0.301, 0.500] & [0.300, 0.500] & [0.500, 0.500] \end{pmatrix}$$

Next, the consistency of the matrix is analyzed to determine if the comments issued by the users have been expressed without ignorance. For this we use the consistency index ICI^{+}:

$$ICI^{+}(M) = 0.989$$

To determine if M it is consistent, it is necessary to select a critical acceptable value for CI^{+} [24]. For this example, we select the 80th percentile (p_{80}), where $\delta_{*} = 0.9930$, then the matrix is not consistent, and we need to apply Algorithm 1. This algorithm consists in systematically decreasing the diameter ($diam(m) = m^{U} - m^{L}$) of the matrix intervals, to find a consistent matrix, if it exists. Algorithm 1 was executed by deflating the intervals with $\Delta = 10^{-5}$. ICI^{+} converged to the appropriate value with this process, it took 9,768 iterations until reaching the limit of the intervals with the consistency value $ICI^{+}(M) = 0.994 \geq \delta_{*} = 0.993$, making the matrix consistent. The resulting matrix is shown in Table 3.

Table 3: Consistent Interval Matrix obtained after a deflation process.

$M^{(9,768)}$ =

Company	MoviStar	Orange	Jazztel	Vodafone
MoviStar	[0.500, 0.500]	[0.397, 0.402]	[0.398, 0.403]	[0.597, 0.602]
Orange	[0.597, 0.602]	[0.500, 0.500]	[0.596, 0.602]	[0.597, 0.600]
Jazztel	[0.596, 0.601]	[0.397, 0.403]	[0.500, 0.500]	[0.597, 0.602]
Vodafone	[0.398, 0.402]	[0.399, 0.402]	[0.397, 0.402]	[0.500, 0.500]

Step 3. Obtaining the Final Alternatives Ranking.

To obtain the final ranking of the alternatives from the interval comparison matrix M (Table 3) we use the AC-OWG operator. The attitude character chosen for this work was $Q(y) = y$, and hence, we have $\lambda = \int_0^1 Q(y)dy = 1/2$. Table 4 shows the telecommunications companies' preference based on their sentiment values.

Table 4: Preference of the alternatives in function of the public's sentiment.

$$AC - OWG(M) = \begin{bmatrix} \psi(MoviStar) \\ \psi(Orange) \\ \psi(Jazztel) \\ \psi(Vodafone) \end{bmatrix} = \begin{bmatrix} 0.9615 \\ 0.9205 \\ 0.7709 \\ 0.4232 \end{bmatrix}$$

This ranking tells us that MoviStar is the company with the most positive feeling in social media in the study period, while Vodafone is the company with the worst public feeling.

5. Discussion and Conclusion

Social media are increasingly present in people's lives and are greatly used for communications. This means that the information generated in these media is increasingly relevant for decision-making processes of companies or organizations. Having decision models that facilitate the contextualization of information, as well as its consistency, avoiding information generated with ignorance, is fundamental to improve those decision processes.

In this work, a model for decision making has been proposed to obtain an alternative ranking contextualizing the feelings/opinions of the users represented through feeling consistent with data extracted from social media. This model uses an interval representation to contextualize the public opinion, showing preferences variation in a more precise way, by incorporating the preferences interval as well as the preferences cardinality. To this end, a majority aggregation operator has been extended to the interval context, bringing about operator ISMA-OWA, which allows the construction of majority preference intervals, based on individual preference valuations, this way building interval pairwise-comparison matrices.

Also, the model incorporates an interval consistency index, ICI^{+}, to measure the consistency of the public preferences, together with a critical value for accepting or rejecting matrices, and thus avoid the use in the decision processes, of data that have been issued with ignorance. In addition, an algorithm for the reconstruction of matrices for the pairwise-comparison has been proposed, which performs a deflation of the preferred interval, searching for a consistent data subset. The model also proposes an aggregation operator, to obtain the priorities and hence the ranking of the alternatives ranking from interval pairwise-comparison matrices with additive relations.

Finally, the model has been tested in a real case, with data from the Spain's telecommunications sector. For this purpose, information has been taken from social media. The results obtained are consistent with what was expected based on the real communications expressed in social media.

As future work we intend to continue with the automation of the decision processes. To achieve this goal, one of the problems to be solved is to determine the decision criteria that decision-makers use in reality. To do this, we will use the model proposed in this work, which allows to more accurately contextualize the data and its consistency.

References

[1] Y. Dong, G. Zhang, W.C. Hong, Y. Xu, Consensus models for AHP group decision making under row geometric mean prioritization method, Decis. Support Syst. 49 (2010) 281–289. doi:10.1016/j.dss.2010.03.003.

[2] N. Capuano, F. Chiclana, H. Fujita, E. Herrera-Viedma, V. Loia, Fuzzy Group Decision Making with Incomplete Information Guided by Social Influence, IEEE Trans. Fuzzy Syst. (2017) 1–1. doi:10.1109/TFUZZ.2017.2744605.

[3] J. Bernabé-Moreno, A. Tejeda-Lorente, C. Porcel, H. Fujita, E. Herrera-Viedma, Quantifying the emotional impact of events on locations with social media, Knowledge-Based Syst. 146 (2018) 44–57. doi:10.1016/J.KNOSYS.2018.01.029.

[4] D.C. . Zikopoulos, K. Parasuraman, T. Deutsch, J. Giles, Harness the Power. Of Big Data The IBM Big Data Platform, McGraw Hill Professional, New York, NY, 2012.

[5] K. J. Trainor, M. T. Krush, R. Agnihotri, Effects of relational proclivity and marketing intelligence on new product development, Mark. Intell. Plan. 31 (2013) 788–806. doi:10.1108/MIP-02-2013-0028.

[6] P. Ross, C. McGowan, L. Styger, A comparison of theory and practice in market intelligence gathering for Australian micro-businesses and SMEs, Soc. Sci. Res. Netw. (2012) 1–17. http://www.ssrn.com.

[7] J. Leskovec, Social Media Analytics : Tracking , Modeling and Predicting the Flow of Information through Networks, World Wide Web Internet Web Inf. Syst. (2011) 277–278. doi:10.1145/1963192.1963309.

[8] L. Dey, S.M. Haque, A. Khurdiya, G. Shroff, Acquiring competitive intelligence from social media, Proc. 2011 Jt. Work. Multiling. OCR Anal. Noisy Unstructured Text Data - MOCR_AND '11. (2011) 1. doi:10.1145/2034617.2034621.

[9] H. Zhang, Y. Dong, E. Herrera-Viedma, Consensus Building for the Heterogeneous Large-Scale GDM with the Individual Concerns and Satisfactions, IEEE Trans. Fuzzy Syst. 26 (2018) 884–898. doi:10.1109/TFUZZ.2017.2697403.

[10] R. Ureña, F. Chiclana, H. Fujita, E. Herrera-Viedma, Confidence-consistency driven group decision making approach with incomplete reciprocal intuitionistic preference relations, Knowledge-Based Syst. 89 (2015) 86–96. doi:10.1016/J.KNOSYS.2015.06.020.

[11] T.L. Saaty, L.G. Vargas, Uncertainty and rank order in the analytic hierarchy process, Eur. J. Oper. Res. 32 (1987) 107–117. doi:10.1016/0377-2217(87)90275-X.

[12] R. Collobert, J. Weston, A unified architecture for natural language processing, Proc. 25th Int. Conf. Mach. Learn. - ICML '08. (2008) 160–167. doi:10.1145/1390156.1390177.

[13] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, P. Kuksa, Natural Language Processing (Almost) from Scratch, J. Mach. Learn. Res. 12 (2011) 2493–2537. doi:10.1.1.231.4614.

[14] J. Wu, F. Chiclana, E. Herrera-Viedma, Trust based consensus model for social network in an incomplete linguistic information context, Appl. Soft Comput. 35 (2015) 827–839. doi:10.1016/J.ASOC.2015.02.023.

[15] B. Dehe, D. Bamford, Development, test and comparison of two Multiple Criteria Decision Analysis (MCDA) models: A case of healthcare infrastructure location, Expert Syst. Appl. 42 (2015) 6717–6727. doi:10.1016/J.ESWA.2015.04.059.

[16] K. Govindan, M.B. Jepsen, ELECTRE: A comprehensive literature review on methodologies and applications, Eur. J. Oper. Res. 250 (2016) 1–29. doi:10.1016/j.ejor.2015.07.019.

[17] M. Marttunen, J. Lienert, V. Belton, Structuring problems for Multi-Criteria Decision Analysis in practice: A literature review of method combinations, Eur. J. Oper. Res. 263 (2017) 1–17. doi:10.1016/J.EJOR.2017.04.041.

[18] T.L. Saaty, The Analytic Hierarchy Process, Education. (1980) 1–11. doi:10.3414/ME10-01-0028.

[19] J. Buchanan, P. Sheppard, Ranking Projects Using the ELECTRE Method, Oper. Res. Soc. New Zealand, Proc. 33rd Annu. Conf. (1998) 42–51.

[20] J. Brans, P. Vincke, A preference ranking organization method: the PROMETHEE method for MCDM, Manage. Sci. 31 (1985) 647–656.

[21] R.E. Moore, R.B. Kearfott, M.J. Cloud, Introduction to Interval Analysis, 2009. doi:10.1137/1.9780898717716.

[22] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hil, New York, 1980. doi:10.3414/ME10-01-0028.

[23] M. Hu, B. Liu, Mining and summarizing customer reviews, Proc. 2004 ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. - KDD '04. (2004) 168. doi:10.1145/1014052.1014073.

[24] J.I. Peláez, E.A. Martínez, L.G. Vargas, Consistency in Positive Reciprocal Matrices: an Improvement in Measurement Methods, IEEE Access. (2018). doi:https://doi.org/10.1109/ACCESS.2018.2829024.

[25] J.I. Peláez, R. Bernal, M. Karanik, Majority OWA operator for opinion rating in social media, Soft Comput. 20 (2016) 1047–1055. doi:10.1007/s00500-014-1564-6.

[26] J.I. Peláez, J.M. Doña, Majority multiplicative ordered weighting geometric operators and their use in the aggregation of multiplicative preference relations., Mathw. Soft Comput. 12 (2005) 107–120. http://eudml.org/doc/40862.

[27] M. Karanik, J.I. Peláez, R. Bernal, Selective majority additive ordered weighting averaging operator, Eur. J. Oper. Res. 250 (2016) 816–826. doi:10.1016/j.ejor.2015.10.011.

[28] R.R. Yager, Z. Xu, The continuous ordered weighted geometric operator and its application to decision making, Fuzzy Sets Syst. 157 (2006) 1393–1402. doi:10.1016/j.fss.2005.12.001.

[29] R.R. Yager, On ordered weighted averaging aggregation operators in multi criteria decision making, IEEE Trans. Syst. Man Cybern. 18 (1988) 183–190. doi:10.1109/21.87068.

[30] R.R. Yager, Quantifier guided aggregation using OWA operators, Int. J. Intell. Syst. 11 (1996) 49–73. doi:10.1002/(SICI)1098-111X(199601)11:1<49::AID-INT3>3.3.CO;2-L.

[31] R. Weisman, An Overview of TOGAF Version 9.1, Publ. by Open Gr. (2011) 43.

[32] J.I. Peláez, J.M. Doña, Majority additive-ordered weighting averaging: A new neat ordered weighting averaging operator based on the majority process, Int. J. Intell. Syst. 18 (2003) 469–481. doi:10.1002/int.10096.

[33] F. Liu, Acceptable consistency analysis of interval reciprocal comparison matrices, Fuzzy Sets Syst. 160 (2009) 2686–2700. doi:10.1016/j.fss.2009.01.010.

[34] J.I. Peláez, M.T. Lamata, A new measure of consistency

for positive reciprocal matrices, Comput. Math. with Appl. 46 (2003) 1839–1845. doi:10.1016/S0898-1221(03)90240-9.
[35] R.R. Yager, OWA aggregation over a continuous interval argument with applications to decision making, IEEE Trans. Syst. Man, Cybern. Part B Cybern. 34 (2004) 1952–1963. doi:10.1109/TSMCB.2004.831154.
[36] W. He, S. Zha, L. Li, Social media competitive analysis and text mining: A case study in the pizza industry, Int. J. Inf. Manage. 33 (2013) 464–472. doi:10.1016/j.ijinfomgt.2013.01.001.
[37] W. He, H. Wu, G. Yan, V. Akula, J. Shen, A novel social media competitive analytics framework with sentiment benchmarks, Inf. Manag. 52 (2015) 801–812. doi:10.1016/j.im.2015.04.006.
[38] E. Cambria, B. Schuller, Y. Xia, C. Havasi, New Avenues in Opinion Mining and Sentiment Analysis, IEEE Intell. Syst. 28 (2013) 15–21. doi:10.1109/MIS.2013.30.
[39] D.R. Tobergte, S. Curtis, NLP with Python, 2013. doi:10.1017/CBO9781107415324.004.
[40] J. Yang, Z. Liu, C. Jia, K. Lin, Z. Cheng, New Data Publishing Framework in the Big Data Environments, in: 2014 Ninth Int. Conf. P2P, Parallel, Grid, Cloud Internet Comput., IEEE, 2014: pp. 363–366. doi:10.1109/3PGCIC.2014.139.
[41] M. Chau, Hsinchun Chen, Research feature - Comparison of three vertical search spiders, Computer (Long. Beach. Calif). 36 (2003) 56–62. doi:10.1109/MC.2003.1198237.
[42] T.L. Saaty, A scaling method for priorities in hierarchical structures, J. Math. Psychol. 15 (1977) 234–281. doi:10.1016/0022-2496(77)90033-5.