Promoting consensus in the concept mapping methodology: an application in the hospitality sector

Albert Fornells[a], Zaida Rodrigo[a], Xari Rovira[c], Mónica Sánchez[d], Ricard Santomà[a], Francesc Teixidó-Navarro[a,b], Elisabet Golobardes[b]

[a]*Research Group in Hospitality, Tourism and Mobilities, School of Tourism and Hospitality Management Sant Ignasi, Universitat Ramon Llull. Av. Marquès de Mulhacén 40-42, 08034 Barcelona, Spain*
[b]*GR-SETAD, La Salle, Universitat Ramon Llull. Av. Quatre Camins 30, 08022 Barcelona, Spain*
[c]*ESADE Business School, Universitat Ramon Llull. Av. Pedralbes 62, 08034 Barcelona, Spain*
[d]*Universitat Politècnica de Catalunya. UPC-Barcelona Tech., Jordi Girona, 1-3, 08034 Barcelona, Spain*

ABSTRACT

The concept mapping methodology aims to respond to the non trivial task of conceptualising abstract thoughts by means of a focus group composed by experts from the studied domain. The approach defines a set of general steps that allow experts to lead the generation of ideas, group the ideas in a conceptual map of interrelated concepts using clustering multidimensional scaling and clustering techniques, analysing the quality of the conceptual maps and deciding on a final interpretation. In this sense, this final decision is not trivial because clustering techniques provide a set of potentially conceptual maps so experts must select the one that fits best according to their opinion. For this reason, we present the global index of consensus as an indicator for filtering the most suitable clustering solutions using qualitative reasoning. It promotes the consensus of experts opinions and ensures objectivity in the final interpretation. The index outperforms three of the most well-known clustering validation indexes in a case study focused on the meaning of excellence in hospitality industry. This work presents the global index of consensus as an indicator for filtering the most suitable clustering solutions using qualitative reasoning that promotes the consensus of experts' opinions, which is one of the key aspects in the concept mapping methodology. The index outperforms three of the most well-known clustering validation indexes in a case study focused on the meaning of excellence in hospitality.

1. Introduction

The concept mapping methodology aims to respond the challenge of guiding a group of experts in the objective representation of thoughts, ideas or abstract concepts based on promoting their agreement regarding what they consider more relevant in consensus (Trochim, 1989; Bigné et al., 2002). Thus, this method is used to offer clarity and develop a model or specify a conceptual framework and it has been successfully applied in education, social research and management science fields to create conceptual frameworks based on specific aspects (Nabitz et al., 2001). The methodology defines a set of general steps using qualitative and quantitative data to determine a conceptual map of interrelated concepts (Rosas and Camphausen, 2007). Giving a specific topic study through a set question, a focus group composed of experts in this domain generate ideas related to this topic using brainstorming. Next, the focus group have to group and weight the ideas in categories based on their point of view. This information is converted into knowledge using data mining techniques (Witten and Frank, 2011), which are applied to identify shared patterns between the opinion of the experts using multidimensional scaling and clustering techniques. It is important to highlight that clustering techniques often return more than one possible solution where each one represents a clustering configuration that groups elements in a specific way. Therefore, the last step is to validate and select the most suitable clustering configuration based on the criteria of the group of experts. Although one of the main benefits of this approach is its flexibility and adaptability, the amount of data that has to be analysed may hinder the tasks of experts because the selection of the best clustering configuration is non trivial and they have to review all the results following the sub-

jective premise "does it make sense to you?" Trochim (1989), which may compromise the objectivity of the approach.

This paper presents the global index of consensus (GIc) to help experts in selecting the most suitable clustering configuration based on two of the main premises of the concept mapping methodology: objectivity and consensus. Thus, the knowledge discovery process is drastically improved because experts have to focus only in useful configurations characterised for containing ideas in which the experts agree are similar and with the same relevance. This index is based on qualitative reasoning techniques and the concept of entropy (Shannon, 1948). Qualitative reasoning is a sub area of artificial intelligence that seeks to understand and explain human beings' non numerical evaluations and it also permits to handle with non numerical data preserving the principle of relevance, i.e., each variable can be valued with the level of precision required (Travé-Massuyès et al., 2004, 2005). Finally, GIc is successfully evaluated and compared with respect to other approaches for tackling one of the challenges of the tourism sector: 'what are the main factors that lead to excellence in hospitality?'

This article is structured as follows. Section 2 describes the concept mapping methodology and how quantitative validation index can be used for selecting the most suitable patterns. Section 3 proposes the global index of consensus and describes its bases. Section 4 applies the concept mapping methodology for discovering the meaning of excellence in hospitality, and it also show how this index outperforms the results provided by some of the most well-known quantitative index. Finally, section 5 ends with the conclusions and further work.

2. Framework

This section summarises the concept mapping methodology and some of the most well-known validation indexes used for the selection of the most suitable clustering solution.

2.1. Concept Mapping methodology

Concept mapping was developed by Trochim (1989) to respond to the conceptualisation needs based on the objectification of opinions and ideas from a group of experts. It uses a methodology which incorporates statistical techniques, such as multidimensional scaling and cluster analyses, and its applications is based on six main steps as indicated in Figure. 1: preparation, generation of statements, structuring statements, representation of statements, interpretation of maps and utilisation of maps.

1. **Preparation**. The aim of the preparation step is twofold. On one hand, clarify the construct for research. On the other, choose the focus group members who will participate in the process. For the best results, the group should contain up of 8-15 participants who are as diverse as possible in order to have different points of view reflected (Harrison and Klein, 2007).

2. **Generation of Statements**. Upon finishing the first step, the participants are invited to offer their ideas regarding on the main topic at hand using a brainstorming session. The development of this phase usually counts on the collaboration of an expert team specialised in group dynamics in order to obtain the best results possible (Bigné et al., 2002; Calvo et al., 2006).

3. **Structuring of Statements**. The purpose of this step consists in determining how the different statements raised in the previous process are related. For this, the participants are asked to, firstly, evaluate each of the statements mentioned during the brainstorming session and, secondly, each member had to group the different statements according to their own criteria. For the first part, the focus group members are asked to rate the list of statements using a Likert scale from 1 to 5 (Bigné et al., 2002) according to the degree of adjustment of the statement towards the concept set out. For the second part, participants have to group the list items according to their own criteria and a label that represents the main concept of each group they considered. Once the participants have rated the items and grouped them, a similarity matrix S_{rxr} for each of the participants is created where r represents the total number of statements generated during the brainstorming session. In each intersection, a '1' was introduced if a person in the group had put both items in the same group; otherwise, a '0' was introduced. Next, each individual's matrix is added to create a general grouping matrix as shown in Figure 1. In the central diagonal there is the total number of participants, and the number of each intersection shows the number of people who put both corresponding statements in the same group independently of their meaning or the criteria used (Calvo et al., 2006).

4. **Representation of Statements**. Two statistical techniques are applied to objectify the results obtained during the previous phase. The data gathered in the grouping matrix in the previous step is processed using a multidimensional scaling (MDS) technique as agreed by Borg and Groenen (1997) to project the original r dimensional space into a two dimensional space where each dot represents an statement. The closer together two dots are, the greater the number of people is who feel that these statements were in the same group. Once the map is constructed, a clustering algorithm (Jain, 2010) is applied using the same distance coordinates as those obtained by the MDS (Borg and Groenen, 1997). Clustering algorithms are able to group data from different points of view, and their suitability mainly depends on the application domain. Thus, the elements in a cluster are similar among them and different from the elements of other clusters and it provides the experts a possible classification or categorisation of the elements.

One of the main challenges that experts have to tackle when they apply clustering techniques is the selection of the most suitable cluster solution as these techniques usually offer more than one possible solution, and there is no exact method to determine the definitive number of clusters as stated by Hair et al. (2006). Thus, the expert needs to evaluate the different configurations generated by the approach in order to select the most suitable representation of the statements.

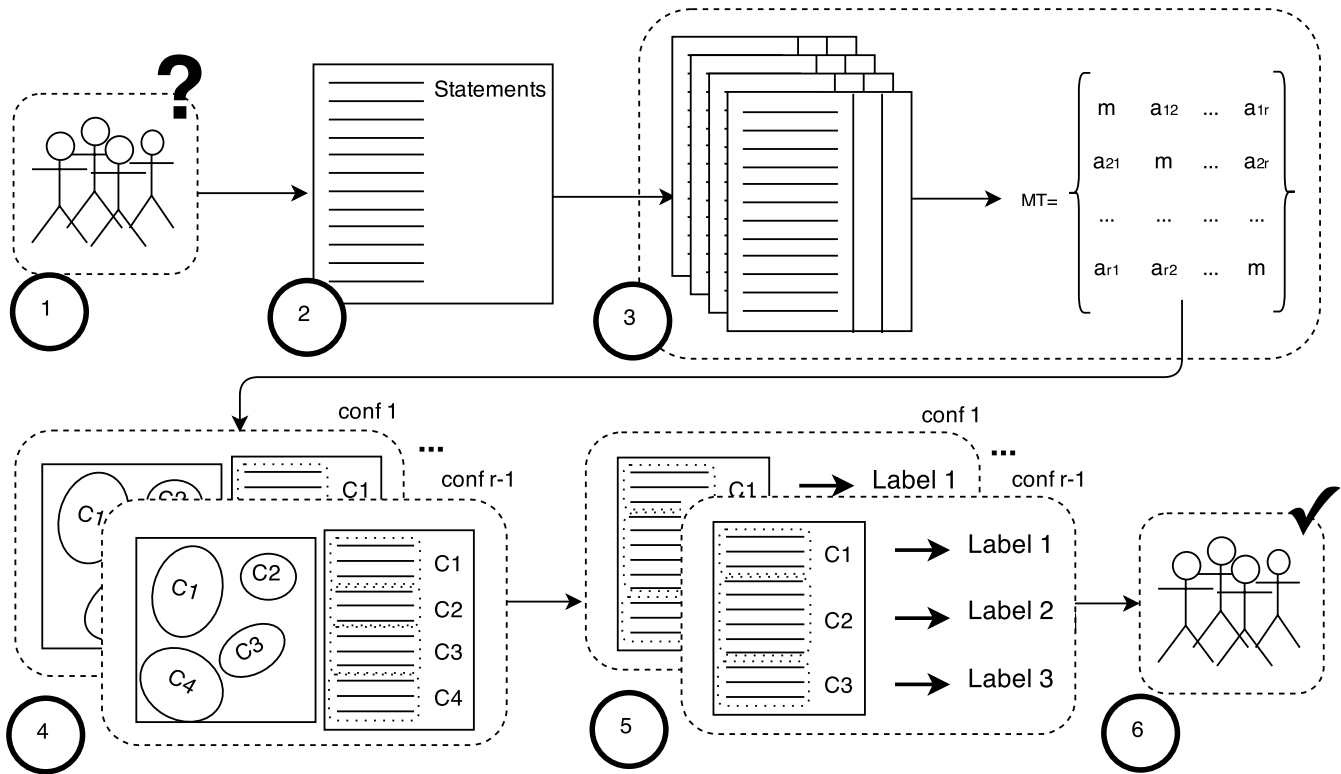For this reason, the usage of a cluster validation index has

Statements

$$MT = \begin{Bmatrix} m & a_{12} & ... & a_{1r} \\ a_{21} & m & ... & a_{2r} \\ ... & ... & ... & ... \\ a_{r1} & a_{r2} & ... & m \end{Bmatrix}$$

1

2

3

conf 1

conf r-1

C1

C1

C2

C1

C2

C3

C4

conf 1

conf r-1

C1 → Label 1

C1 → Label 1

C2 → Label 2

C3 → Label 3

4

5

6

Fig. 1. The concept mapping process is split into 6 steps: **(1)** A set of experts are selected for finding out the meaning of a specific concept; **(2)** A list of r ideas is generated through a Brainstorming process; **(3)** Ideas are evaluated by the experts and this information is used to build a matrix $MT_{r \times r}$; **(4)** A multi dimensional scaling and a clustering technique are applied over the matrix to project the information in a 2D space. The result is a set of r-**1** possible cluster configurations; **(5)** Experts analyse all the configurations and they label each one of the clusters based on their items; **(5)** Experts agree on selecting the best concept representation based on the subset of concepts identified **(6)**.

become crucial to facilitate the data analysis in order to score or sort the possible solutions based on indicators (also called validation index) that promotes the separability and/or the compactness of the clusters (Garcia-Piquer et al., 2012, 2014). These indicators allow experts to reduce the range of potentially valid solutions because they are based only on quantitative values without taking into account information related to the problem, which limits their capabilities. In this sense, the definition of a specific index based on the domain characteristics is the best way to help experts to select the most suitable solution (Corral et al., 2011). This is exactly what we do in this article by means of the GIc, which allows experts to rank the clustering solutions based on what experts agree.

5. **Interpretation of Maps**. After the statements had been mapped via the cluster analysis, a name is given to represent the key statements in each group as shown in Figure 1. To carry out this step, interviews are held with experts in the field to evaluate the content of each group. The participating experts analyse the obtained results so that the global result corresponds to reality (Trochim, 1989).

6. **Utilisation of maps**. Lastly, the maps are used as a graphic representation of the experts' opinion regarding the concept under study (Bigné et al., 2002).

One of the biggest benefits of this methodology is its flexibility and capability for being applied to any kind of domain because it defines a framework that can be easily adapted to fit the problem that has to be faced. In contrast, this flexibility becomes a challenge because there is no one single way of doing things and this may hamper analysis. A clear example occurs in step 4 where the application of the clustering technique offers a set of potentially valid configurations. In that case, the selection of the configuration will set the difference between the success or failure and that is where the expert can be aided to select the most suitable configuration using the key element: consensus.

2.2. Looking for the most suitable patterns

The use of unsupervised learning approaches such as clustering algorithms is an essential step in almost any data analysis problem. However, two independent steps are needed before hand: analysts have to select (1) the clustering approach and (2) select the most appropriate solution from the whole set of possible solutions.

There is not a single criteria to classify the clustering algorithms, so they can be classified according to many criteria (Gan et al., 2000; Witten and Frank, 2011; Duda et al., 2000): (1) the search strategy to find the clusters (center-based, graph-based, model-based, search-based, density-based and subspace clustering), (2) the relationships between the clusters (partitional and hierarchical), (3) the instances distribution into the clusters

(hard clustering or fuzzy clustering), and (4) the optimisation of the clusters (conventional clustering, ensemble clustering, or multiobjective clustering). For this reason, it is important to select the algorithm according to the data typology and the features of the application domain (Jain, 2010).

On the other hand, clustering techniques usually offer more than one possible cluster solution because they are not able to automatically identify the optimal number of groups to discover or the results are conditioned by some initial parameters that need to be tuned among other reasons (Garcia-Piquer, 2012). These facts often hinders the data analysis step because experts must evaluate all the different solutions generated by the algorithm, which is highly time consuming and quite arbitrary because the selection will depend on the subjectivity of the expert due to the fact that all of them are potentially valid. For this reason, the application of evaluation functions for automatically scoring the clustering solutions has become the key for helping experts to select the best (Gurrutxaga et al., 2011). These evaluation functions define metrics that measure the cluster quality by using the same features included in the data set. Therefore, the challenge is define what quality means as the following subsections describe.

2.2.1. Deviation and Connectivity

The simplest evaluation metric for measuring the quality of a cluster solution is to evaluate how close the elements of each cluster are and how separated the clusters between them are. Consequently, the smaller their values the better is the solution. This is exactly what Deviation and Connectivity measures do respectively (Handl and Knowles, 2007; Hruschka et al., 2009; Gurrutxaga et al., 2011).

The *Deviation* (Dev) measures the compactness of the clusters. It is computed as the overall summed distances between data items and their corresponding cluster centre as Equation 1 shows, where C is the clustering obtained, C_i is the set of instances belonging to cluster i, v_i is the centroid of cluster i, and $d(x, v_i)$ is the Euclidean distance between the element x and v_i.

The *Connectivity* (Conn) refers to the cluster connectedness. It takes into account the degree to which data points that are close in the feature space have been placed in the same cluster as Equation 2 shows, where r is the number of examples in the training data set, C is the clustering obtained, $nn(x, i)$ returns the ith nearest element of x using the Euclidean distance and ℓ is the amount of nearest elements taken into account. Note that, for each instance i, the metric computes a weighted sum of the ℓ nearest neighbours that belong to a different cluster from that of i (the weight is decreased according to how far instances i and j are).

Although the information provided by the Deviation and the Connectivity allows to obtain insights from the cluster solution analysed, this information is not enough to select the best configuration. Thus, it is necessary to define more powerful indicators to evaluate the clustering solutions.

$$Dev(C) = \sum_{i \in C} \sum_{x \in C_i} d(x, v_i) \tag{1}$$

$$Conn(C) = \sum_{x=1}^{r} \left(\sum_{i=1}^{\ell} \chi(x, nn(x, i), i) \right), \text{ where} \tag{2}$$

$$\chi(x, y, i) = \begin{cases} \dfrac{1}{i} & \text{if } \neg \exists j : x \in C_j \wedge y \in C_j, \\ 0 & \text{otherwise.} \end{cases}$$

$$DB(C) = \frac{1}{k} \sum_{i=1}^{k} \max_{\substack{j=1 \\ j \neq i}}^{k} \left\{ \frac{S_k(C_i) + S_k(C_j)}{d(v_i, v_j)} \right\} \tag{3}$$

$$S_k(C_i) = \frac{1}{|C_i|} \sum_{x \in C_i} d(x, v_i)$$

$$Dn(C) = \min_{i=1}^{n} \left\{ \min_{j=i+1}^{n} \left\{ \frac{d(C_i, C_j)}{\max_{k=1}^{n} \{diam(C_k)\}} \right\} \right\} \tag{4}$$

$$diam(C_k) = \max_{x,y \in C_k} \{d(x, y)\}$$

$$d(C_i, C_j) = \min_{x \in C_i, y \in C_j} \{d(x, y)\}$$

$$Sil(C) = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{|C_i|} \sum_{x \in C_i} \left(\frac{b(C, i, x) - a(C_i, x)}{max\{a(C_i, x), b(C, i, x)\}} \right) \right)$$

$$a(C_i, x) = \frac{1}{|C_i| - 1} \sum_{y \in C_i} d(x, y) \tag{5}$$

$$b(C, i, x) = \min_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{|C_j|} \sum_{y \in C_j} d(x, y)$$

2.2.2. Cluster validation index

Cluster validation index (Halkidi et al., 2001) is an objective function that evaluates the clustering results. Validation techniques can be based on comparing clusters to the original classes of the problem if classes are known (supervised approach), or by validating clusters according to their quality based on the compacting and separation between them when classes are unknown (unsupervised approach).

Regarding the characteristics of the clustering technique used in the concept mapping (partitional and hard clustering), there are three possible clustering validation approaches (Halkidi et al., 2002; Legány et al., 2006). The reader is referred to (Theodoridis and Koutroumbas, 2008) for hierarchical and fuzzy clustering algorithms. The first one is called external criteria and the idea is to evaluate a clustering result comparing it with a structure of the data set obtained without applying any clustering algorithm. The second approach is called internal criteria and the objective is to evaluate a clustering result comparing it with only quantities and features inherent to the data set. The third approach is called relative criteria and it is based on comparing a clustering result with other results obtained from the application of the same clustering algorithm with different parameter values, or of other clustering algorithms. The cluster validation methods based on external or internal criteria are based on statistical hypothesis testing, and their major drawback is their high computational cost. Moreover, these two approaches measure the degree to which a data set confirms an

a-priori specified scheme that can be inherent to the data set or an intuitive structure of the data. On the other hand, relative criteria methods find the hypothetical best clustering scheme from several clustering results obtained with different parameters or clustering algorithms without using statistical tests, so they are less computationally expensive. Therefore, this last approach is what we really need for our purposes because we want to compare the performance of several configuration in an unsupervised process.

The Davies-Bouldin index (Davies and Bouldin, 1979) (see Eq. 3), the Dunn index (Dunn, 1974) (see Eq. 4) and the Silhouette index (Rousseeuw, 1987) (see Eq. 5) are three of the most well known validation strategies. The range of the first two indexes is $[0, +\infty]$ and $[-1, 1]$ for the last one. Silhouette and Dunn indexes have to be maximised, and Davies-Bouldin index has to be minimised. In the three equations C is the clustering obtained; k is the number of clusters; C_i is the set of instances belonging to cluster i; $|C_i|$ is the number of elements in C_i; v_i is the centroid of C_i; $d(x, y)$ is the Euclidean distance between x and y elements; $nn(x, i)$ returns the ith nearest element of x according to $d(x, y)$; and ℓ is the amount of nearest elements taken into account. The main difference of these three indexes is the calculation of the quality of the shape of each cluster. Davies-Bouldin index evaluates the clusters taking into account if they are scattered, calculating the distance between the instances of each cluster and their respective centroid. Dunn index evaluates the clusters calculating if they are compact, penalising the clusters with a long diameter. Silhouette index calculates the tightness of the clusters, taking into account the distance between the instances of each cluster.

3. Methodological approach

The approach proposed in this paper helps the experts to evaluate in the step 4 the different configurations generated in the previous steps, and to select de optimal clustering solution, as described in Subsection 2.1. To this end, the proposed methodology aims to identify and emphasise the clusters and the ideas in which the experts agree that are the most important ones. This is done in two phases. First a degree of consensus of clusters is defined to measure the agreement among the members of the focus group with respect to each cluster. Second a recurrence index of each idea, highlighting the ideas that are more in consensus, is used as a weight coefficient to define a global index of consensus of configurations in order to promote the ideas in which experts agree. This methodology relies on the use of qualitative labels belonging to a qualitative absolute order-of-magnitude model. This allows dealing with the focus group members evaluations to improve Concept Mapping processes through the new global index.

3.1. Qualitative reasoning

The one-dimensional absolute order-of-magnitude model (Travé-Massuyès et al., 2004, 2005) works with a finite number of qualitative labels corresponding to an ordinal scale of measurement. The number of labels chosen to describe a real problem is not fixed, but depends on the characteristics of each represented variable.

Let us consider an ordered finite set of *basic* labels $\mathbb{S}_* = \{B_1, \ldots, B_n\}$, being n the number of labels chosen each one of them corresponding to a linguistic term, in such a way that $B_1 < \cdots < B_n$.

Example 1. To illustrate the expression of a set of ordered linguistic evaluations, an example of the *basic* labels for $n = 5$ is given by: B_1 = "not important at all" $< B_2$ = "of little importance" $< B_3$ = "important" $< B_4$ = "very important" $< B_5$ = "extremely important".

The complete universe of description for the order-of-magnitude space is the set $\mathbb{S} = \mathbb{S}_* \cup \{[B_i, B_j] \,|\, B_i, B_j \in \mathbb{S}_*, i < j\}$, where the label $[B_i, B_j]$ with $i < j$ is defined as the set $\{B_i, B_{i+1}, \ldots, B_j\}$, with the convention $[B_i, B_i] = \{B_i\} = B_i$.

Consistent with Example 1, the linguistic evaluation "very or extremely important" can be represented by the non-basic qualitative label ["very important", "extremely important"], i.e., $[B_4, B_5]$. The label "unknown" is represented by ["not important at all", "extremely important"], i.e., $[B_1, B_5]$.

The order in the set of basic labels $\mathbb{S}_*$ induces a partial order $\leq$ in $\mathbb{S}$ defined as: $[B_i, B_j] \leq [B_r, B_s] \iff B_i \leq B_r$ and $B_j \leq B_s$. This relation is trivially an order relation in $\mathbb{S}$, but a partial order, since there are pairs of non-comparable labels. For instance, in Example 1, the relation $[B_i, B_j] \leq [B_r, B_s]$ expresses that $[B_i, B_j]$ is "less or equal important than" $[B_r, B_s]$.

There is another partial order relation $\leq_P$ in $\mathbb{S}$ "to be more precise than", given by $[B_i, B_j] \leq_P [B_r, B_s]$ iff $[B_i, B_j] \subset [B_r, B_s]$, i.e. $r \leq i$ and $j \leq s$. The less precise label is $? = [B_1, B_n]$. This structure permits working with all different levels of precision from the basic labels to the $?$ label (see Figure 2).

Two different binary operations are defined in the complete universe of description $\mathbb{S}$, called the connex union and the intersection, introduced in a more general context as the mix and the common operations in Roselló et al. (2010).

Definition 1. Given two qualitative labels $[B_{i_1}, B_{j_1}]$, $[B_{i_2}, B_{j_2}] \in \mathbb{S}$, their *connex union* is the qualitative label $[B_{i_1}, B_{j_1}] \sqcup [B_{i_2}, B_{j_2}] = [B_{\min(i_1, i_2)}, B_{\max(j_1, j_2)}]$.
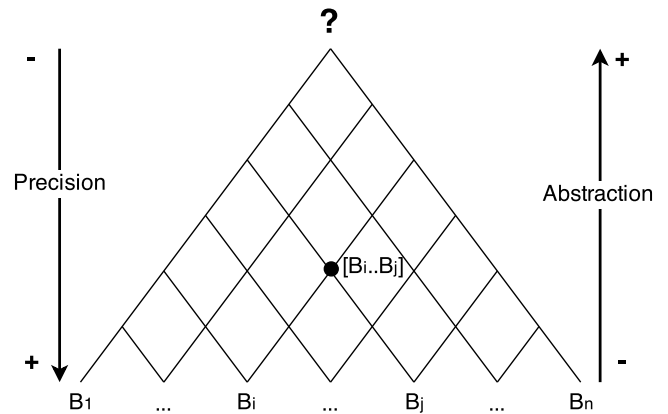

Fig. 2. The complete universe of description $\mathbb{S}$ (Roselló et al. (2010)).

Definition 2. Given two qualitative labels $[B_{i_1}, B_{j_1}]$, $[B_{i_2}, B_{j_2}] \in \mathbb{S}$, such that $[B_{i_1}, B_{j_1}] \cap [B_{i_2}, B_{j_2}] \neq \emptyset$, their *intersection* is the qualitative label $[B_{i_1}, B_{j_1}] \cap [B_{i_2}, B_{j_2}] = [B_{\max(i_1, i_2)}, B_{\min(j_1, j_2)}]$.

Following Example 1, the connex union of B_1 and $[B_3, B_4]$ is $B_1 \sqcup [B_3, B_4] = [B_1, B_4]$, and their intersection is empty. In the case of the pair of labels $[B_1, B_3]$ and $[B_3, B_4]$ their connex union is $[B_1, B_3] \sqcup [B_3, B_4] = [B_1, B_4]$ and their intersection is $[B_1, B_3] \cap [B_3, B_4] = B_3$.

3.2. Entropy and Consensus

A definition of a consensus index, able to synthesise the focus group members' evaluations, is used in the proposed methodology. The consensus index involves the notion of entropy of a qualitative label defined in $\mathbb{S}$, inspired in Shannon entropy concept in information theory (Shannon, 1948). This notion requires a normalised measure on the set $\mathbb{S}$, i.e., a measure such that $\mu(B_i) > 0$ for all $i = 1, \ldots, n$ and $\sum_{B_i \in \mathbb{S}^*} \mu(B_i) = 1$. Then: $\mu([B_i, B_j]) = \sum_{k=i}^{j} \mu(B_k)$, $\mu([B_i, B_j]) \leq 1$ for all $i, j \in \{1, \ldots, n\}$, and $\mu([B_1, B_n]) = 1$.

In the context of qualitative absolute order-of-magnitude models, the concepts of information of a label and entropy of a qualitative description were introduced in Roselló et al. (2010). Following Roselló et al. (2010), in this paper we define the entropy of a qualitative label by a positive continuous real function of the measure of this label as follows:

Definition 3. The *entropy of a qualitative label* $Q \in \mathbb{S}$ is defined as:

$$H(Q) = \ln \frac{1}{\mu(Q)},$$

where μ is the measure considered in $\mathbb{S}$.

Note that for all $Q \in \mathbb{S}$, if $Q \neq ?$ then $\mu(Q) \in (0, 1)$ and, consequently, $H(Q) > 0$. Moreover, H decreases with respect to $\leq_P$: $Q \leq_P Q' \Rightarrow Q \subset Q' \Rightarrow \mu(Q) \leq \mu(Q') \Rightarrow \ln \frac{1}{\mu(Q)} \geq \ln \frac{1}{\mu(Q')}$. In addition, $H(?) = \ln 1 = 0$.

The definition of the degree of consensus of a set of qualitative labels is as follows:

Definition 4. Given m qualitative labels $Q_1, \cdots, Q_m \in \mathbb{S}$, such that $\cap_{j=1}^{m} Q_j \neq \emptyset$, their *degree of consensus* is:

$$Dc(Q_1, \cdots, Q_m) = \frac{H(\sqcup_{j=1}^{m} Q_j)}{H(\cap_{j=1}^{m} Q_j)} = \frac{\ln(\mu(\sqcup_{j=1}^{m} Q_j))}{\ln(\mu(\cap_{j=1}^{m} Q_j))}$$

In the case that $\cap_{j=1}^{m} Q_j = \emptyset$, their *degree of consensus* is $Dc(Q_1, \cdots, Q_m) = 0$.

Example 2. Let us consider the concept C = "Customer Oriented" and a focus group of 3 members $\mathbb{E} = \{e_1, e_2, e_3\}$. Let us assume that the opinions of the three members with respect to C are represented by three qualitative labels defined as: $Q_1(C) = [B_4, B_5], Q_2(C) = B_4, Q_3(C) = [B_3, B_5]$ using the linguistic evaluations corresponding to basic labels $B_1, \ldots, B_5$ given in Example 1. Finally, let us define $\mu(B_i) = 1/5, i = 1, \ldots, 5$.

Then, since $\sqcup_{k=1}^{3}(Q_i(C)) = [B_4, B_5] \sqcup B_4 \sqcup [B_3, B_5] = [B_3, B_5]$ and $\cap_{k=1}^{3}(Q_i(C)) = [B_4, B_5] \cap B_4 \cap [B_3, B_5] = B_4$, the degree of consensus is:

$$Dc(Q_1, Q_2, Q_3) = \frac{H(\sqcup_{k=1}^{3}(Q_i(C)))}{H(\cap_{k=1}^{3}(Q_i(C)))} = \frac{H([B_3, B_5])}{H(B_4)} =$$

$$= \frac{\ln 3/5}{\ln 1/5} = 0.32$$

3.3. The Proposed Global Index of Consensus

The structure of qualitative absolute order-of-magnitude models allows us to deal with the focus group members' evaluations of ideas and concepts in a concept mapping process. To this end, we work in a one-dimensional absolute order-of-magnitude model with n basic labels corresponding to the n ordered responses of the Likert scale used by the members of the focus group.

In the following, let us consider a focus group consisting of m members, that after a brainstorming process have generated and evaluated a set of r different ideas. For each member of the focus group j, with $1 \leq j \leq m$, and for each idea X, the *opinion of member j with respect to X* is an element of $\mathbb{S}_*$, which is denoted by $V_j(X)$.

Note that the concept mapping method provides $r - 1$ different cluster solutions. Each cluster solution provides exactly k clusters or groups, being $1 \leq k < r$. For instance, on the one hand, the cluster solution with 2 clusters groups $r - 1$ ideas in one cluster and the remaining most discordant idea alone in the other cluster. On the other hand, the cluster solution with $r - 1$ clusters, groups the most similar two ideas in a cluster meanwhile the remaining $r - 2$ ideas are each one in a different cluster.

From now on, for each $1 \leq k < r$, clusters belonging to the configuration with k groups will be denoted by C_i^k, with $1 \leq i \leq k$.

Definition 5. Fixed a configuration with k clusters, let j, with $1 \leq j \leq m$, be a member of the focus group, let C_i^k, $1 \leq i \leq k$, be a cluster, and let $\{X_1^i, \ldots, X_{s_i}^i\}$ be the set of ideas in cluster C_i^k. The *opinion of member j with respect to C_i^k* is defined as:

$$Q_j(C_i^k) = V_j(X_1^i) \sqcup \cdots \sqcup V_j(X_{s_i}^i).$$

Note that $Q_j(C_i^k)$ belongs to the complete universe of description for the order-of-magnitude space. In the case that $V_j(X_1^i) = \cdots = V_j(X_{s_i}^i)$, $Q_j(C_i^k)$ is a basic label, otherwise $Q_j(C_i^k)$ is a non-basic label.

Intuitively speaking, $Q_j(C_i^k)$ is the result of mixing the evaluations of member j of all ideas in cluster C_i^k in a new one that includes all of them.

The entropy H and the degree of consensus Dc introduced in Subsection 3.2 allows us to define a measure of consensus among the members of the focus group with respect to each cluster:

Definition 6. Fixed a configuration with k clusters, let C_i^k, $1 \leq i \leq k$, be a cluster, and let $Q_j(C_i^k)$ be the opinion of member j with respect to C_i^k, $1 \leq j \leq m$. The *degree of consensus of cluster C_i^k* is

$$Dc(C_i^k) = \begin{cases} \dfrac{H(\bigsqcup_{j=1}^{m} Q_j(C_i^k))}{H(\bigcap_{j=1}^{m} Q_j(C_i^k))}, & \text{if } \bigcap_{j=1}^{m} Q_j(C_i^k) \neq \emptyset \\ 0, & \text{otherwise.} \end{cases}$$

In order to reflect if a configuration has any cluster with nonzero degree of consensus, we define:

Definition 7. Fixed a configuration with k clusters,

$$N(k) = \begin{cases} 1, & \text{if } \exists C_i^k \text{such that } Dc(C_i^k) \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

The above definition allows us to introduce a recurrence index of an idea. The index is defined via these numbers $N(k)$, k=1 to r-1 and the characteristic functions of clusters. The recurrence index of an idea takes into account the expert's perception of the relevance of the idea, which is extracted from the evaluations of the focus group members.

Definition 8. Let X be an idea, The *recurrence index of X* is:

$$RI(X) = \frac{\sum_{k=1}^{r-1} \sum_{i=1}^{k} 1_{C_i^k}(X)}{\sum_{k=1}^{r-1} N(k)}$$

where $1_{C_i^k}$ is the characteristic function that indicates membership of an element in C_i^k.

In this way, RI(X) provides a normalisation of the number of times the idea X is in consensus among all the configurations.

To be able to select the most suitable cluster solution as an improvement of a concept mapping process, we propose a global index of consensus of each configuration, to measure the cluster solutions quality. The global index of consensus of a configuration is constructed as a weighted sum of the degrees of consensus of clusters in the configuration. Each weight captures the importance of the corresponding cluster by considering the addition of the recurrence index of ideas belonging to it.

Definition 9. The *global index of consensus of a configuration with k clusters* is

$$GIc(k) = \sum_{i=1}^{k} w(C_i^k) \cdot Dc(C_i^k),$$

where $w(C_i^k) = \sum_{X \in C_i^k} RI(X)$, i.e. the addition of the recurrence index of ideas in cluster C_i^k.

4. Case study: What are the main factor that describe excellence in hospitality?

This section tackles the challenge of analysing the meaning of the excellence in hospitality by means of the application of the concept mapping methodology. First, excellence in hospitality is introduced. Next, the application of the first four steps for the statement generation is described. Thus, the selection of the most suitable clustering configuration is analysed by means of the application of quantitative and qualitative index. Finally, experts label and assess the discovered concept through the last two steps.

4.1. Excellence in hospitality

There is an increasing concern over the term and meaning of hospitality as well as a search among academics to identify a globally understood an accepted conceptual term that defines this concept. Slattery et al. Slattery et al. (2002) refers to Lashley et al. (2000) who posit that the understanding of hospitality has been impaired by an industrial myopia and propose to improve the understanding by; reflecting insights into the study of hospitality that encompass the commercial provision of hospitality and the hospitality industry, yet at the same time it is recognised that hospitality needs to be explored in a private domestic setting and studies the concept of hospitality as a social phenomenon involving relationships between people. Along similar lines of thought, Wood and Brotherton (Wood and Brotherton, 2008) affirm the statement that the conceptual development for the hospitality concept is limited, and the academic literature that does exist is scattered. This current issue and underlying challenge to find the meaning of hospitality can be rooted by the essence of the discipline. The hospitality sector has always been a professional one and from a commercial and management point of view it has evolved into more areas where an interaction between a host and a guest takes place. Even though a holistic approach to the term hospitality is increasingly accepted, it is still important to define the key aspects that define such discipline not only for academic purposes but also to help develop better suited professionals (Vila et al., 2012). Hospitality has always been about relationships. The word itself means friendliness to strangers. Hemmington (2007) suggests that by exploring and defining hospitality as an experience, new perspectives emerge that have important implications for hospitality in commercial contexts. These implications take hospitality beyond services management to a place where hospitableness, a sense of theatre and generosity are central. Hemmington advocates that hospitality businesses must focus on the guest experience and stage memorable experiences that stimulate all five senses. By achieving this, hospitality organisations that are able to capture this sense of theatre and generosity will gain competitive advantage by providing their guests with experiences that are personal, memorable and add value to their lives. Linking this concept to the values in the hospitality concept, (Harkison et al., 2011) posit that hospitality, because of its connection to values and ways of thinking is not an art, a science, or a business, but a philosophy. It further proposes that those motivated by a desire to serve are strongly and deeply attracted to work in commercial hospitality where they can express themselves

and find meaning through their work. In those places where there is an interaction between hosts and guests, loyalty is the critical factor for a sustainable business. To achieve loyalty in hospitality, an attitude and philosophy towards excellence is necessary. Business excellence is important in creating sustainable and continuous quality improvement of business processes, that may bring strong financial performance, high customer demand, goal achievement, successful employee recruitment and admission, desired product and service outcome, and outstanding staff (Kanji, 2002).

4.2. Generation of the possible maps

The steps 1, 2, 3 and 4 of the concept mapping methodology were done in the following way for creating all the set of potentially groups of concepts.

1. **Preparation**. A focus group composed by 11 experts representing the hospitality industry in Barcelona was held in January 2014. All participants were senior managers with more than 10 years of relevant experience in international companies. It is important to highlight that Barcelona is the 10th-most-visited city in the world and the third most visited in Europe after London and Paris, with 8 million tourists every year since 2012 (Bremmer, 2011). Barcelona is a internationally renowned tourist destination with numerous recreational areas, historical monuments, including eight UNESCO World Heritage Sites, many good-quality hotels, and developed tourist infrastructure. The participants were asked to answer the question: From your point of view, what are the main factors that describe excellence in hospitality?

2. **Generation of Statements**. A list of 100 ideas were generated following the brainstorming session leaded by an expert in group dynamics (see Table 1 in the supplementary material).

3. **Structuring of Statements**. The aggregation matrix was built by means of the addition of the matrix of each focus group member. Each individual's matrix represents the evaluation of ideas from 1 to 5 and how the ideas are related among them.

4. **Representation of Statements**. The multidimensional scaling (in accordance with the Alscal Method) was applied to project the information in two dimensions. Next, a clustering method (according to the Ward Method) was executed to generate the potentially statement representations that were from only 1 cluster ($k=1$) since 99 clusters ($k=99$).

4.3. Selection of the most suitability map representation

Steps 5 and 6 are focused on labelling and analysing the correctness of the most suitable clustering configuration from all the r potentially configurations discovered in the step 4 and, consequently, this decision will condition the results. We analyse this selection from three different perspectives using: (1) Quantitative indexes, (2) The proposed global index of consensus and (3) Combining the mentioned quantitative indexes and the consensus concept, using the degree of consensus (Dc) introduced in Definition 6.

Figure 3 shows the results of applying the Davis-Bouldin, Dunn and Silhouette indexes described in Subsection 2.2.2 which are only based on the cluster geometric data without taking into account additional knowledge from the domain. The figure shows that Davis-Bouldin and Dunn indexes obtain the best results when the k is higher. In contrast, Silhoutte index promote the solutions with fewer clusters. Therefore, all index are promoting extreme solutions that will hinder the task of experts because the concept has to be defined using few clusters with almost all ideas or many clusters with one or two ideas. For this reason, quantitative indexes are not enough for guiding the selection process. If the expert is looking for concepts that should represent the consensus between the different focus group members, why do not introduce this domain characteristic into the problem?

The application of the global index of consensus requires three steps: (1) Preprocess the 99 clustering configurations in order to remove the clusters and elements without consensus (Definition 4); (2) Calculate the recurrence index (Definition 8) as Figure 5 shows; (3) Apply the global index of consensus (Definition 9) over the preprocessed clustering configurations as Subsection 3.3 describes. Figure 4 shows the global index of consensus and the impact of applying the consensus concept over the clusters and their elements into the 99 configurations. For each one of the configurations, the clusters (and their items) without consensus are removed because they represent information that introduces noise and uncertainty. As seen, the number of clusters and elements is drastically reduced because this preprocessing operation promotes the ideas in which experts agree. Thus, the configuration $k=20$ is selected as the best because it has the highest global index of consensus and the highest number of clusters and elements in consensus.

Finally, Figure 6 shows the application of the Davis-Bouldin, Dunn and Silhouette indexes described in Subsection 2.2.2 over the 99 preprocessed configurations. The most significant effect that can be observed from the application of that consensus concept is that the linearity of the scoring is broken because some empty configurations are rejected, which is reducing the original scope of configurations. Nevertheless, each index promotes different configurations as it happened before and $k=85$ is selected as the best cluster configuration based on the three indexes because it is the intersection between them.

4.4. Analysis and assessment of the discovered concept

The best configurations selected for the global index of consensus and for the combination of the consensus concept with the quantitative indexes were presented to the experts in order to label the clusters and assess its meaning regarding to the initial question. After the analysis of both configurations, the experts agreed that the configuration selected using the quantitative indexes was too small for extracting any conclusion because it had only two clusters and seven ideas in consensus. In contrast, the configuration provided by the proposed global index of consensus allowed them to label the resulting clusters as:

- Innovation, quality, successful service performance and technology assistance ideas represent **Service Quality**. An
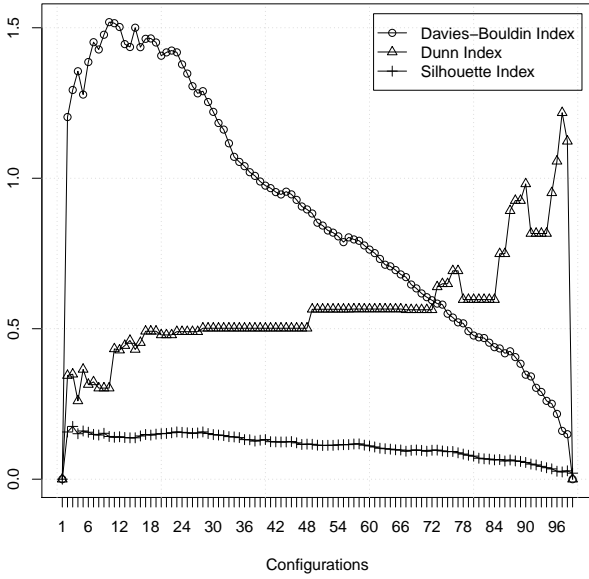
Fig. 3. The Davis-Bouldin, Dunn and Silhouette indexes are applied over the 99 configurations generated in the step 4.Silhouette and Dunn indexes have to be maximised, and Davies-Bouldin index has to be minimised. They are optimising only the cluster geometry without taking into account the most key element: the consensus of the experts.
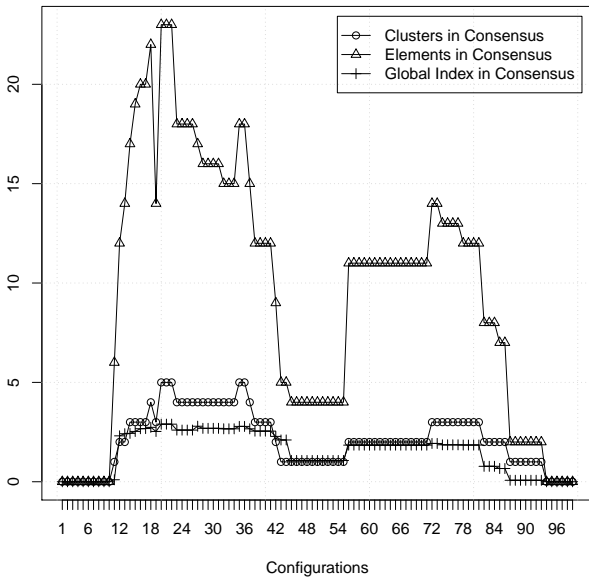

Fig. 5. Ranking of the ideas based on a recurrence index greater than zero for each one of the 99 configurations.

Fig. 4. The graphic shows the global index of consensus and how many clusters (and how many items) have a degree of consensus different from zero for each one of the 99 configurations. This approach highlights the most relevant and useful data, which are the ones with consensus. Thus, the global index of consensus is a qualitative index that promotes the selection of configurations based on the elements with highest degree of consensus.
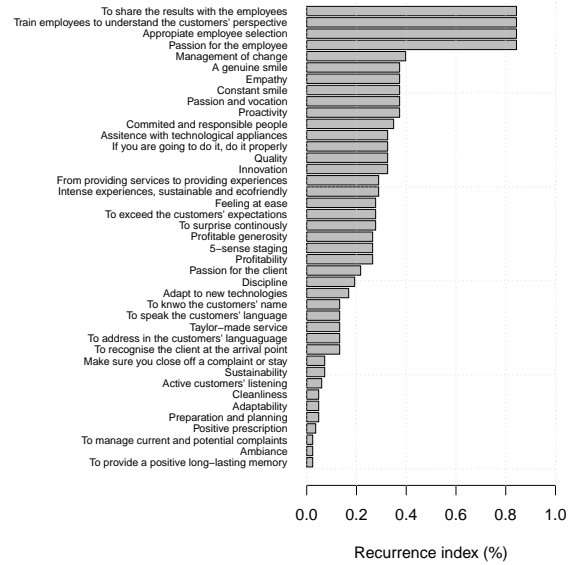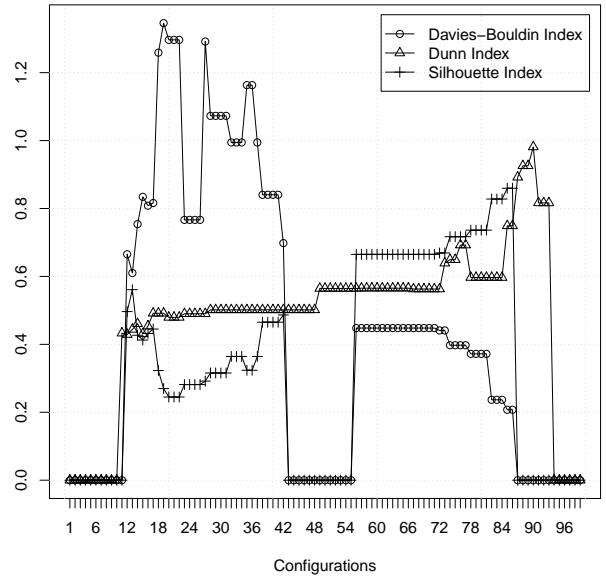

Fig. 6. The Davis-Bouldin, Dunn and Silhouette indexes are applied over the 99 preprocessed configurations generated in the step 4. The preprocess consists in removing the clusters (and its elements) with degree of consensus equal to zero. The value of index is 0 for those configurations that are too small and it is not possible to compute the index values.

excellent hospitality means to deliver an innovative and excellent service to the customer.

- To acknowledge the customer upon arrival, To address

to the customer in his/her language, tailor-made service, speak the customer in his or her language and to identify the customer by his or her name represents **Customer Oriented**. Description: An excellent hospitality means to have a company, which understands and fulfils customer needs and expectations.

- To surprise continuously, exceeding the customers expectations, feel at home and positive prescription represents **Loyalty**. An excellent hospitality means to create a long-lasting relationship with the customer by exceeding his or her expectations and making the customer feel special.

- Passion for the employee, change management, discipline, an appropriate employee selection, to train the employees to understand the customers needs and to share the quality results with the employees represents **Human Resource management**. An excellent hospitality is delivered by excellent people at the firm, so an excellent human resource management is the key to achieve it.

- 5 senses setting, an intense, sustainable and eco-friendly experiences, to adapt to new trends and to shift from service delivery to experience creation represents **Creating an experience**. An excellent hospitality means providing memorable experiences rather than a plain service.

Finally, we also reviewed the worst scenarios to figure out why they scored with a low value. In this sense, experts analysed the configurations k=2, 3 and 11 selected by the Silhoutte, Dunn and Davies-Bouldin respectively and their conclusion was that they were not useful. In fact, Figure 4 shows that the consensus for these configurations is 0.

5. Conclusion and further work

One the biggest challenges when applying the concept mapping methodology is to decide the suitability or the relevance of a configuration. In this work a new method that substantially improves the clustering determination in this methodology has been proposed. A global index of consensus is defined based on the objectivity and the consensus, which are two of the main premises of the concept mapping methodology. The proposed global index helps experts to select the most suitable clustering configuration improving the knowledge discovery process, since experts have to focus only on useful configurations. In that sense, the global index helps experts to identify the configurations containing concepts and ideas in which experts agree for being the most important ones. The global index presented in this paper is based on qualitative reasoning techniques and permits extract valuable and useful information from experts, which is crucial to select the most suitable configuration.

The proposed methodology has been applied to analyse the meaning of excellence in hospitality in a case study framed in the Barcelona hospitality industry in January 2014. The case study experimental results proved that this method achieves much better results compared with current state-of-the-art approaches, based only on quantitative data.

From a theoretical point of view, future work includes the adaptation of the proposed global index to other unsupervised clustering methodologies to determine the most suitable configuration. On the other hand, understanding what excellence in hospitality means from a cross-cultural perspective should be very interesting because the mobility in this century is a key element in tourism.

References

Bigné, J.E., Manzano, J.A., Küster, I., Vila, N., 2002. The concept mapping approach in marketing: an application in the travel agencies sector. Qualitative Market Research: An International Journal 5, 87–95.

Borg, I., Groenen, P., 1997. Modern multidimensional scaling. New York: Springer.

Bremmer, C., 2011. Euromonitor internationals top city destinations ranking. Euromonitor International .

Calvo, A., Criado, F., Periez, R., 2006. Desarrollo de un instrumento para evaluar la idoneidad de los planes docentes: una aplicacin a la diplomatura en turismo. Presented in Decisiones basadas en el conocimiento y en el papel social de la empresa. Academia Europea de Direccin y Economa de la Empresa, Palma de Mallorca.

Corral, G., Garcia-Piquer, A., Orriols-Puig, A., Fornells, A., Golobardes, E., 2011. Analysis of vulnerability assessment results based on {CAOS}. Applied Soft Computing 11, 4321 – 4331. Soft Computing for Information System Security.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 1, 224–227.

Duda, R., Hart, P., Stork, D., 2000. Pattern classification. New York, John Wiley and Sons, Inc.

Dunn, J., 1974. Well separated clusters and optimal fuzzy partitions, in: Journal of Cybernetics, pp. 95–104.

Gan, G., Chaoqun, M., Wu, J., 2000. Data clustering theory, algorithms, and applications. ASA-SIAM, Philadelphia.

Garcia-Piquer, A., 2012. Facing-up challenges of multiobjective clustering based on evolutionary algorithms: Representations, scalability and retrieval solutions. Ph.D. thesis. Research Group in Intelligent Systems, Campus LaSalle, Universitat Ramon Llull.

Garcia-Piquer, A., Fornells, A., Bacardit, J., Orriols-Puig, A., Golobardes, E., 2014. Large-scale experimental evaluation of cluster representations for multiobjective evolutionary clustering. IEEE Transactions on Evolutionary Computation 18, 36 – 53.

Garcia-Piquer, A., Fornells, A., Orriols-Puig, A., Corral, G., Golobardes, E., 2012. Data classification through an evolutionary approach based on multiple criteria. Knowledge and Information Systems 33, 35–56.

Gurrutxaga, I., Muguerza, J., Arbelaitz, O., Prez, J.M., Martn, J.I., 2011. Towards a standard methodology to evaluate internal cluster validity indices. Pattern Recognition Letters 32, 505 – 515.

Hair, J.F., Tatham, R.L., Anderson, R.E., Black, W., 2006. Multivariate Data Analysis. 6th ed., Prentice Hall.

Halkidi, M., Batistakis, Y., Vazirgiannis, M., 2001. On clustering validation techniques. Journal of Intelligent Information Systems 17, 107–145.

Halkidi, M., Batistakis, Y., Vazirgiannis, M., 2002. Cluster validity methods: Part i. SIGMOD Record 31, 40–45.

Handl, J., Knowles, J., 2007. An evolutionary approach to multiobjective clustering. IEEE Transactions on Evolutionary Computation 1, 56–76.

Harkison, T., Poulston, J., Kim, J.H.G., 2011. Hospitality graduates and managers: the big divide. International Journal of Contemporary Hospitality Management 23, 377–392.

Harrison, D.A., Klein, K.J., 2007. What's the difference? diversity constructs as separation, variety, or disparity in organizations. Academy of management review 32, 1199–1228.

Hemmington, N., 2007. From wervice to experience: Understanding and defining the hospitality business. The Service Industries Journal 27, 747–755.

Hruschka, E.R., Campello, R.J.G.B., Freitas, A.A., de Carvalho, A.C.P.L.F., 2009. A survey of evolutionary algorithms for clustering. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews 39, 133–155.

Jain, A.K., 2010. Data clustering: 50 years beyond k-means. Pattern recognition letters 31, 651–666.

Kanji, G., 2002. Measuring business excellence. Routledge advantages in management and business studies .

Lashley, C., Morrison, A., et al., 2000. In search of hospitality: theoretical perspectives and debates. In search of hospitality: theoretical perspectives and debates. .

Legány, C., Juhász, S., Babos, A., 2006. Cluster validity measurement techniques, in: Proceedings of the 5th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, pp. 388–393.

Nabitz, U., Severens, P., Brink, W.V.D., Jansen, P., 2001. Improving the efqm model: An empirical study on model development and theory building using concept mapping. Total Quality Management 12, 69–81.

Rosas, S.R., Camphausen, L.C., 2007. The use of concept mapping for scale development and validation in evaluation. Evaluation and Program Planning 30, 125 – 135.

Roselló, L., Prats, F., Agell, N., Sánchez, M., 2010. Measuring consensus in group decisions by means of qualitative reasoning. International Journal of Approximate Reasoning 51, 441 – 452.

Rousseeuw, P., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, in: Journal of Computational and Applied Mathematics, pp. 53–65.

Shannon, C., 1948. A mathematical theory of communication. Bell System Technical Journal 27, 379–423, 623–656.

Slattery, P., et al., 2002. Finding the hospitality industry. Journal of Hospitality, Leisure, Sport and Tourism Education 1, 19–28.

Theodoridis, S., Koutroumbas, K., 2008. Pattern recognition. 4 ed., Academic Press, Burlington, USA.

Travé-Massuyès, L., Ironi, L., Dague, P., 2004. Mathematical foundations of qualitative reasoning. AI Magazine 24, 91–106.

Travé-Massuyès, L., Prats, F., Sánchez, M., Agell, N., 2005. Relative and absolute order-of-magnitude models unified. Annals of Mathematics and Artificial Intelligence 45, 323–341. doi:10.1007/s10472-005-9002-1.

Trochim, W.M., 1989. An introduction to concept mapping for planning and evaluation. Evaluation and Program Planning 12, 1 – 16. Special Issue: Concept Mapping for Evaluation and Planning.

Vila, M., X., R., Costa, G., Santom, R., 2012. Combining research techniques to improve quality service in hospitality. Qual Quant , 795–812.

Witten, I.H., Frank, E., 2011. Data mining: Practical machine learning tools and techniques with java implementations. 3th ed., Morgan Kaufmann, San Francisco.

Wood, R., Brotherton, B., 2008. The SAGE Handbook of Hospitality Management. SAGE Publications.