Information Sciences 328 (2016) 418–434

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Creating knowledge databases for storing and sharing people knowledge automatically using group decision making and fuzzy ontologies

J.A. Morente-Molinera [a,*], I.J. Pérez [b], M.R. Ureña [a], E. Herrera-Viedma [a]

[a] Department of Computer Science and Artificial Intelligence, University of Granada, Granada, Spain
[b] Department of Computer Sciences and Engineering, University of Cádiz, Cádiz, Spain

ARTICLE INFO

Article history:
Received 22 January 2015
Revised 16 June 2015
Accepted 31 August 2015
Available online 9 September 2015

Keywords:
Fuzzy linguistic modeling
Group decision making
Computing with words
Multi-granular linguistic information
Fuzzy ontology

ABSTRACT

Over the last decade, the Internet has undergone a profound change. Thanks to Web 2.0 technologies, the Internet has become a platform where everybody can participate and provide their own personal information and experiences. Ontologies were designed in an effort to sort and categorize all sorts of information. In this paper, an automatized method for retrieving the subjective Internet users information and creating ontologies is described. Thanks to this method, it is possible to automatically create knowledge databases using the common knowledge of a large amount of people. Using these databases, anybody can consult and benefit from the retrieved information. Group decision making methods are used to extract users information and fuzzy ontologies are employed to store the collected knowledge.

1. Introduction

Originally, the Internet was designed as a means of consulting. Only a few recognized experts were able to provide information while the rest of the people, who could afford to access it, were only allowed to carry out consulting tasks. Therefore, the Internet was designed for minorities and the information available, as compared to what is available nowadays, was quite limited. Nowadays, the situation has changed dramatically. Thanks to Web 2.0 [5,53], the Internet has become a place where users can connect and share large amounts of information. Therefore, Internet users have become providing and consuming information entities. This situation has made information more accessible and available than ever. Nevertheless, in most cases, the information available is badly structured and, therefore, of little use for users. Users just cannot manage all the available amount of information by themselves. In order to deal with this problem, fields like Big Data [39,46,49,51], for extracting conclusions from the data, semantic web [1,9,21,36,47], for sorting it, and intelligent systems [28,38,54], to use the information for different purposes, have arisen.

Ontologies [6,29,43] are tools that provide a way of sorting, classifying and describing large amounts of information. Knowledge databases created using ontologies are easy to manage and allow users to search for information and extract conclusions. Because our system needs to work with conceptual information provided by users, imprecision must be dealt with. For this reason, fuzzy ontologies [17] will be used. Crisp ontologies allow each element to be described or not, {0, 1}, by each concept in the

* Corresponding author. Tel.: +34 958241773.
E-mail addresses: jamoren@decsai.ugr.es (J.A. Morente-Molinera), ignaciojavier.perez@uca.es (I.J. Pérez), raquel@decsai.ugr.es (M.R. Ureña), viedma@decsai.ugr.es (E. Herrera-Viedma).

http://dx.doi.org/10.1016/j.ins.2015.08.051
0020-0255/

ontology. On the contrary, fuzzy ontologies associate each element to each concept using a particular degree in the interval [0, 1]. This way, elements can be associated to different concepts. For example, when referring to a person's height, if John measures 1.78 m, it can be stated in a fuzzy ontology that John has a medium height of 0.7 degree and a high height of 0.3 degree. In a crisp ontology, those measures can be represented as a medium height of 1 and a high height of 0 or vice versa. Therefore, it is easy to see that a fuzzy ontology has a more flexible representation capability than a crisp one.

Retrieving information from Internet users is a quite complicated task. Especially, when subjective information is being dealt with. The provided data must be analyzed and verified. In order to carry out this task, Group Decision Making (GDM) methods [16,58,59] can be used. They allow a set of Internet users to provide information, carry out debates and make a final choice. If this approach is used, the final information outcome is not an outlier opinion from a unique user but a consensus opinion totally guaranteed by a majority of users that have dealt with the matter. Consequently, the obtained information can be considered reliable.

As previously mentioned, the information under consideration was provided by humans. Consequently, it is quite important to provide users with tools that will assist them in providing information. The easier the method of providing information, the more reliable the obtained data will be. For this reason, multi-granular linguistic information [33,50,52] is used in this paper. As a result, Internet users can express themselves using words instead of numbers, i.e. they will be able to express themselves as they are more used to doing so. Moreover, users will be able to select the linguistic term set (LTS) that they want to use to express themselves. This way, if the users want to be very concrete about the provided information, they can use a LTS with a high granularity, otherwise they can select an LTS with a lower granularity.

In this paper, the design of an automatic process for creating knowledge databases using people's common knowledge about a certain issue is being presented. GDM methods are used in order to obtain information totally guaranteed by the majority of users. In order to ease the way in which Internet users provide their opinions, multi-granular linguistic information is used. Finally, the trusted information is automatically stored in a fuzzy ontology where other users can benefit from the obtained knowledge and therefore reach conclusions.

Thanks to the automatic method designed, users can share their subjective knowledge about a certain topic and allow other people to take advantage of it. Retrieved information is sorted in a fuzzy ontology allowing a complete exploitation of the available data. Subjective information provided by human beings is difficult to deal with due to the fact that it is difficult to measure and validate. Thanks to our system, a tool for dealing with this type of information is presented. Moreover, the information used is validated and made objective because it is ratified by the majority of users in the GDM process. In such a way, stored information is no longer an individual person's opinion. It is actually the opinion of a majority, information worthy of being used and taken into account.

In Section 2, some concepts that are required to understand the designed process are presented. In Section 3, the current state of the art on ontologies and decision making applications is presented. In Section 4, the designed process structure is described. In Section 5, an example is given. In Section 6, we make a comparison between the state of the art and our own proposal by analyzing its advantages and drawbacks. Finally, some conclusions are offered.

2. Preliminaries

In order to make this paper as self-contained as possible, this section will introduce concepts and methods to be referred to throughout this paper. In 2.1, we explain how GDM methods work. In Section 2.2, how to deal with multi-granular linguistic information is shown. In 2.3, fuzzy ontologies are described.

2.1. Group decision making

Group Decision Making has been a well-studied field since its first appearance in the 80's [10] until today [15,27,55,68,69]. It has been used satisfactorily in fields such as operations research [70,79], politics [37,62], social psychology [23,42], artificial intelligence [14,18] and soft computing [56,57].

A group decision making problem is formally defined as follows. Let $E = \{e_1, \ldots, e_n\}$ be a set of experts and $X = \{x_1, \ldots, x_m\}$ a set of alternatives. A GDM problem tries to sort X using the preferences values P^k, $\forall k \in [1, n]$, provided by the experts. Generally, GDM processes are carried out following these steps:

1. **Providing preferences**: Experts provide their preferences for the set of alternatives. Thus, in this paper, the next three methods could be used:
   - *Utility functions:* Consist of providing a score value to each of the alternatives. Generally it is considered as a set $U^k = \{u_1^k, \ldots, u_m^k\}$ where $u_\cdot^k \in [0, 1]$ and k is a specific expert.
   - *Preference orderings:* Experts are asked to sort the alternatives according to their preferences. Formally, each expert provides a set $O^k = \{o^k(1), \ldots, o^k(m)\}$ where $o^k(\cdot)$ is a permutation function over the indexes $\{1, \ldots, m\}$.
   - *Preference relations:* Experts provide a pairwise comparison of all the possible pairs of alternatives. Formally, a fuzzy preference relation is a matrix $P \subset X \times X$ where each value is defined by a membership function $\mu_{pk} : X \times X \to [0, 1]$. $\mu_{pk}(x_i, x_j) = p_{ij}^k$ indicates the preferences of x_i over x_j.

   One disadvantage of utility functions is that they do not allow experts to establish relations among the alternatives, information is only based on absolute scoring. Preference orderings establish relations among the alternatives but, information