Chapter 15
Towards a New Extracting and Querying Approach of Fuzzy Summaries

Ines Benali-Sougui
University of Tunis El Manar, Tunisia

Minyar Sassi Hidri
Imam Abdulrahman Bin Faisal University, Saudi Arabia

Amel Grissa-Touzi
University of Tunis El Manar, Tunisia

ABSTRACT
Diversification of DB applications highlighted the limitations of relational database management system (RDBMS) particularly on the modeling plan. In fact, in the real world, we are increasingly faced with the situation where applications need to handle imprecise data and to offer a flexible querying to their users. Several theoretical solutions have been proposed. However, the impact of this work in practice remained negligible with the exception of a few research prototypes based on the formal model GEFRED. In this chapter, the authors propose a new approach for exploitation of fuzzy relational databases (FRDB) described by the model GEFRED. This approach consists of 1) a new technique for extracting summary fuzzy data, Fuzzy SAINTETIQ, based on the classification of fuzzy data and formal concepts analysis; 2) an approach of assessing flexible queries in the context of FDB based on the set of fuzzy summaries generated by our fuzzy SAINTETIQ system; 3) an approach of repairing and substituting unanswered query.

INTRODUCTION
Techniques based on data summaries are now considered as a good way to handle large amounts of data, especially when the precise values of these data are not needed. Thus, several databases (DB) summarization model have been proposed such as SaintEtiQ model (Raschia, 2002), which is close to our
research tasks. This system makes it possible to generate a hierarchy of summaries making it possible to cover parts of the database.

Unfortunately, all of these techniques are limited to the exact data and cannot be applied to fuzzy data. We are confronted more and more with the situation where applications need to manage fuzzy data and to make profit their users from flexible querying. We speak then about flexible querying and Fuzzy Databases (FDB) (Galindo, J., M. P., & Urrutia, A., 2006; Ben Hassine, M.A., Grissa-Touzi, A., Galindo, J. & Ounelli, H., 2008).

Among the results of the query, the user is more interested in the most important or the best, called the Top-K answers. The main reason for this interest is that they avoid overloading the user with a large number of uninteresting answers.

The querying mechanism also provides the means to detect the failed queries reasons. The purpose of the repairing approaches is to provide satisfactory answers even when the query does not accept a result in the strict sense. However, existing approaches only handle Boolean queries and cannot be applied to FRDB.

In this work, we present an extension of the SaintEtiQ summarization model for modeling fuzzy data and therefore to exploit to flexible FSQL query by turning the top \(k\) result. Our goal is also to found a result even in the case of the absence of summaries corresponding strictly to the query.

The rest of the paper is organized as follows: Section 2 presents the basic concepts of Fuzzy Database (FDB), fuzzy attributes in GEFRED model and the basic concept of fuzzy FCA. Section 3 exposes the fuzzy queries’ modeling and an overview of FCA-based summary model. Section 4 gives the limits of the existing summarization approach and of the exploitation of fuzzy data. Then it presents the architecture of our proposed approach. Section 5 describes the first step data organization of our proposed approach. Section 6 describes the second step the querying phase of our proposed approach. Section 7 gives a new kind of repaired query. Section 8 concludes the paper and gives some future work.

BACKGROUND

In this section, we present the basic concepts of FDB, Fuzzy attributes in GEFRED model and the theoretical foundations of fuzzy FCA.

Fuzzy Databases (FDB)

In this section, we present the basic concepts of FDB. A FDB is an extension of the relational database. This extension introduces fuzzy predicates under shapes of linguistic expressions that, at the time of a flexible querying, permits to have a range of answers (each one with a membership degree) in order to offer to the user all intermediate variations between the completely satisfactory answers and those completely dissatisfactory (Bosc, P., O., & Litard. L., 1998). The FRDB models are considered in a very simple shape and consist in adding a degree, usually in the interval \([0,1]\), to every tuple. It allows maintaining the homogeneity of the data in DB. The main models are those of PradeTestemale (Prade, H., & C. T., 1987), Umano-Fukami (Umano, M., M. M. K. T., & Fukami, S., 1980), Buckles-Petry (Buckles, B. P. & F. E. P., 1982), ZemankovaKaendel (Zemankova-Leech, 1985) and GEFRED of Medina et al. (1994). This last model constitutes an eclectic synthesis of the various models published so far with the aim of dealing with the problem of representation and treatment of fuzzy information by using relational DB.