Fuzzy Approach in Enquiry to Regional Data sources for Municipalities

STANISLAVA SIMONOVA, JAN CAPEK
Institute of System Engineering and Informatics
Faculty of Economics and Administration
University of Pardubice
Studentská 84, 532 10 Pardubice,
CZECH REPUBLIC

Abstract: Data represent is an important element of public administration. There are data for basic evidence registers of state administration, data important for the management of self-administration or data for the actual work of institutions. Data are descriptive and geographical, and are used both by institutions of public administration and regional organizations, and of course by citizens. A great problem of the informational environment of public administration is to put a proper enquiry, so that the system would be able to realize the given enquiry both in terms of holding data base and correct enquiry interpretation. The solution is a preparation of a suitable informational environment. Thus, to get a quality answer to an inaccurately entered user enquiry, a certain pre-processing is necessary; when a fuzzy approach, which enables handling vague terms of natural language with the help of mathematical methods, seems to be suitable.

Key-words: Data sources, regional data warehouses, enquiry, fuzzy approach,

1. Introduction

Information systems in public administration institutions and their regional data sources are supposed to serve for analysing and answering strategic regional enquiries, so that this informational environment would serve for an effective area administration and also as a information resource for other institutions and for citizens. A public administration informational source shows certain specificity, by which it differs from informational sources of other organizations [1] [2]. Enterprise informational sources (understood except institutions of public administration) have a program build-up aimed directly at its users. Output mechanisms are controlled by specialists in the given subject, i.e. employees with knowledge in professional terminology and with required level of technological knowledge on the level of their profession. In enterprise data warehouse, it is a "duty" of the user himself to master all necessary operations. Thus, it is not necessary to create a special user interface, which would take into account the different level of knowledge at the enquiry mechanism entry.

Information sources within public administration are used by users with various structures of their professions and with different knowledge in information technologies, or more precisely with unfamiliarity.

2. Data Sources for Municipalities with the Population to 5.000 Citizens from Czech Republic point of view

Region administration is significantly related to the creation of an effective informational environment for the support of this administration; whether the region is understood as the area of a district or a different area [5]. This is expressed and supported by approved documents, instructions and laws within the Czech Republic and the EU.

Table 1 shows that most of municipalities within the Czech Republic belongs to the area which has less than 5.000 populations. From that reason is very important to deal with problem of user friendly information technologies. Municipalities mostly have not ICT specialisation workplace both enough of finance sources or impossibility obtaining qualified workers.

Tab.1: **No. of municipalities [%] from number of inhabitants point of view**

Municipalities in Czech Republic	**No.of municipalities, %**
Municipalities to 4999 Population	95,8%
Municipalities up 5000 to 9999 Population	2,1%
Municipalities up 10000 Population	2,1%

3. Data warehousing

Information systems within public administration are most frequently realized by database software. Above all, these are so-called transaction database systems, which are designed for work with operational data of the organization. Transaction systems work with actual operational data, they are, however, less suitable for analyses in time relations, more complex enquiries etc. Another approach to data sources is brought in by data warehouse technologies, when data are drawn from heterogeneous sources of transaction applications and are stored for a certain period of time, so that they can be used for comparisons, analyses and predictions [4].

Regional data warehouse can be realized on different levels of elaboration and efficiency. The simplest architecture is represented by the application of enquiry tools directly on operational data; the second is the model of local data marts, which already involves separation of enquiry mechanism from operational data and creation of subject-oriented enquiries; the third possible solution is represented by independent data marts, extended by the process of extraction and transformation of heterogeneous data; and finally, the most ideal, complicated and financially demanding is the architecture of enterprise data warehouse with dependent data marts. The selection of a suitable variant of data warehouse architecture must reflect current needs and possibilities of regional institutions.

3.1 Fuzzy approach in user-enquiry

An important function of regional data warehouse is providing quality outputs on even more complicated or so-called inaccurately formulated enquiries [3]. Enquiry creators are represented by public administration personnel, staff of other organizations and especially citizens, i.e. enquiries are put with different accuracy, different knowledge in information technologies and different knowledge in public administration problems. Public administration is, among others, obliged to preserve social cohesion, i.e. it cannot exclude generations or groups that cannot use the technologies from the right to information. The solution is a preparation of a suitable informational environment. Thus, to get a quality answer to an inaccurately entered user enquiry, a certain pre-processing is necessary; when a fuzzy approach, which enables handling vague terms of natural language with the help of mathematical methods, seems to be suitable.

As mentioned above most of members of public administration management are unable to create qualified enquiry.

Regional data sources of information systems should enable the realization of not only an enquiry prepared within the software support, but also of a newly defined enquiry, whereas the level of definition is different (according to the varied composition of data environment users). When the specificity of enquiries about regional data sources is set, it is possible to trace the aspects of the method of enquiry input, enquiry creator, enquiry quality, and search with the help of key words.

Enquiry creator:

- *an expert in both public administration problems and information technologies*: such an employee will not have any problems with the user enquiry input,
- *an expert in public administration*: public administration employee only with limited knowledge in information technologies (for the needs of his profession): the familiarity with terminology enables a good orientation in the user menu, but the creation of a new enquiry represents a problem,
- *a non-expert*: citizens or employees of other organization with minimal knowledge in the public administration professional terminology and with various levels of knowledge in information technologies.

Enquiry input method:

- *prepared enquiry*: by the choice of a suitable option within the menu; the output is prepared (in Figure 1),
- *new enquiry*: either by writing the relevant enquiry command in the enquiry language or by the choice of key words within the search system (in Figure 2).

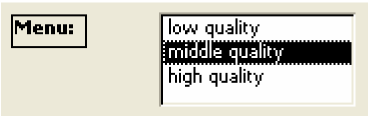


Figure 1: *Prepared enquiry for "Middle quality of some regional coefficient".*

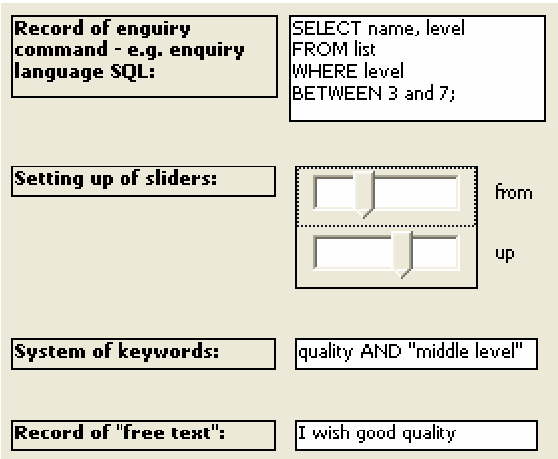


Figure 2: Non-p*repared enquiries for "Middle quality of some regional coefficient".*

Search with the help of key words:

- *inquirer is not familiar with the given professional terminology*: e.g. a citizen wants to get certain data from the resident register (open to him by the law), but the term register is not known to him and he tries to find or enter key words instead, e.g. list, statement, agenda, database etc.,
- *certain object characteristics can have different name in different professional or civil groups*: it is a professional terminology or jargon: e.g. terms like residential unit, structural unit, region, municipality etc.;
- *inquirer tries to enter an enquiry within the area of interest*, which is absolutely strange to him; here, it is not only about the unfamiliarity with terminology, but the area of interest as a whole is little familiar to him, still, he needs to get certain information.

Enquiry quality:

- *targeted enquiry*: correctly and effectively formulated by an expert in information technologies; enquiry is created in the enquiry language or by search in the system of links trough key words (in Figure 3),
- *inaccurately specified enquiry*: caused by technological or professional ignorance; these are either errors during the creation of a new enquiry or the inquirer, due to an inaccurate familiarity with professional terminology, does not orientate himself well in the user menu.

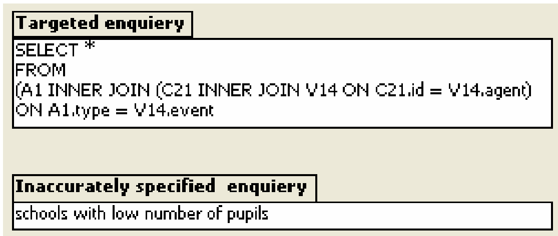


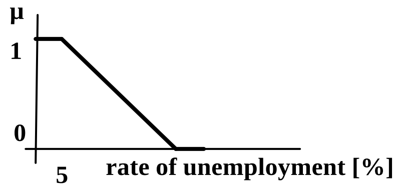
Figure 3: *Enquiry quality*

3.1.1 Specification of membership function

There is necessary to have (create) "general membership function for various regional coefficients (social, economic etc.).

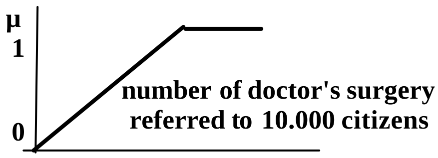
Tab.2: Specification of membership functions for coefficients - employment, medical care and primary education

Employment - membership function:



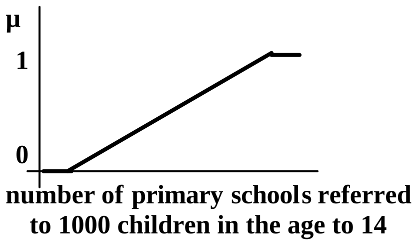
- rate of unemployment to 5% is held as positive, an absolute employment - μ=1
- rate of unemployment over 20% is held as negative - degree of membership μ=0

Medical care - membership function:



- negative medical care (without any doctor's surgery) means - μ=0
- positive medical care (from certain number of doctor's surgery) - μ=1

Primary education - membership function:

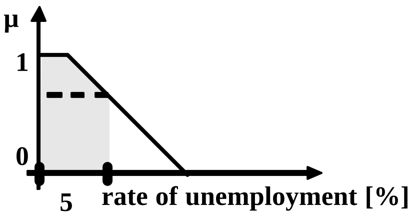


- number of primary schools to certain value is held as low and quite unsatisfactory, it means μ=0
- number of primary schools to certain value is held as excessive and unnecessary, it means degree of membership. μ=1

Table 2 shows membership functions for selected "social" indicators - employment, medical care and primary education.
Then we can evaluate region according to its "social conditions, it means we search crisp number from composed membership functions.

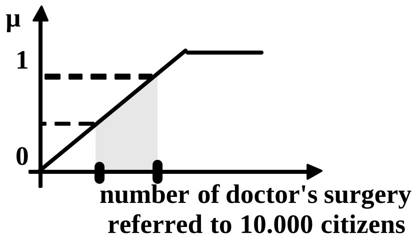
Tab.3: Getting back to crisp number from composed membership functions

"convenient" employment:



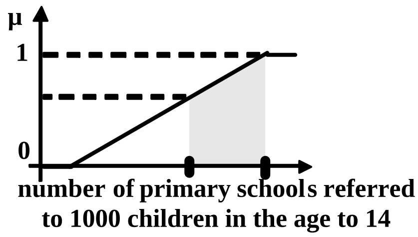
- μ=<0,6; 1>
- corresponding interval of degrees of membership

"convenient " medical care:



- μ=<0,4; 0,7>
- corresponding doctor's surgery numbers (referred to number of population)

"convenient " primary education:



- μ=<0,6; 1>
- corresponding numbers of primary schools

On the figure 4 is shown the chart flow of the user enquiry by means of using fuzzy approach.

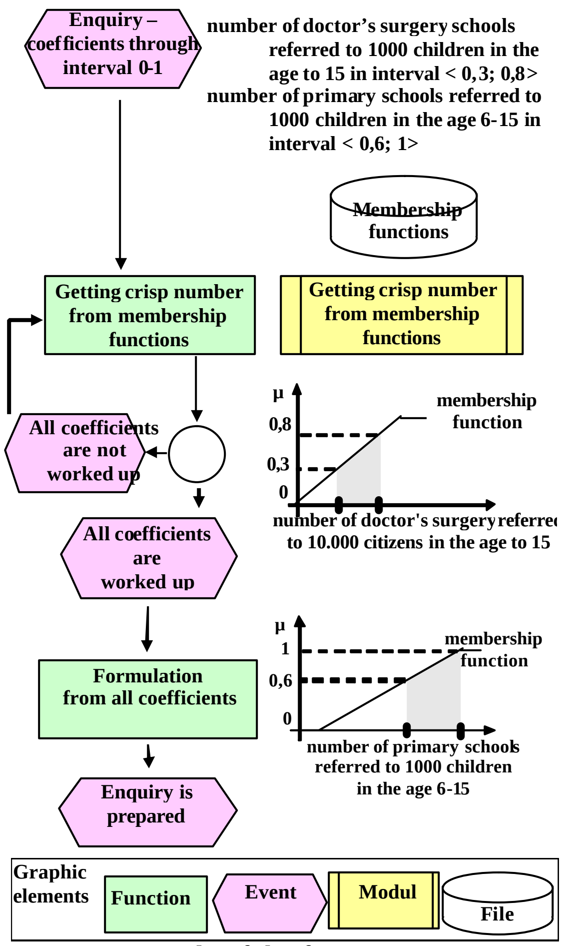


Figure 4: Example of the fuzzy enquire

3.1.2 Unknown words processing

It is obvious from the summary, that an enquiry whose realization is included in the output part of the regional information system, is basically trouble free, because provides only such a form of output, for which it was designed / programmed. To get information with the help of newly created enquiries, certain skills are required, whether these are basic knowledge (e.g. SQL enquiry language) or experience with searching with the aid of key words.

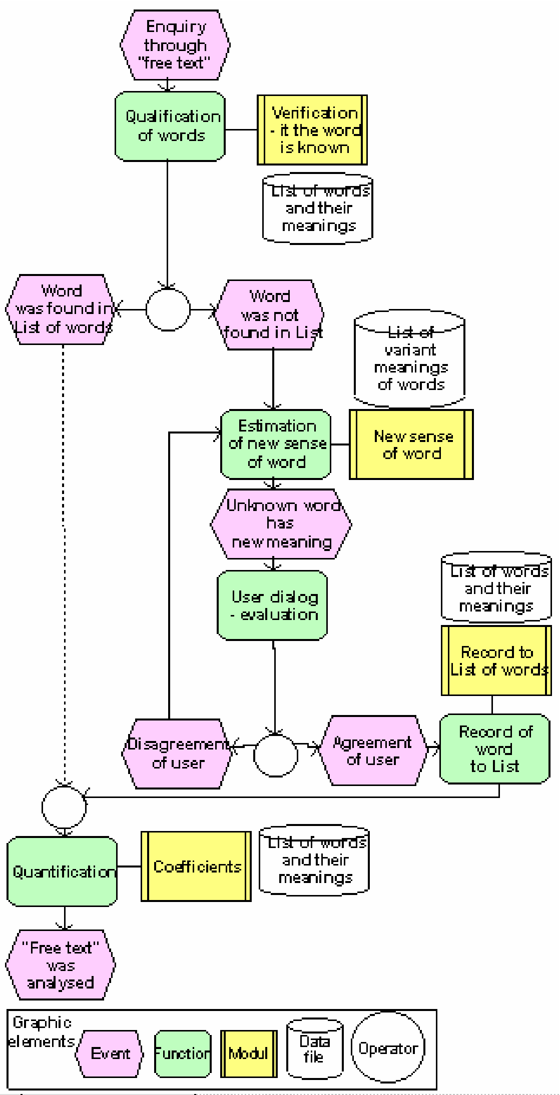


Figure 5: *Analysis of inaccurately specified enquiry*

An expert in the area of information and communication technologies, especially when he/she is an expert in public administration at the same time, he/she is able to create targeted enquiries, which are correctly and efficiently formulated. An interest, however, must be aimed at non-experts in the area of information technologies, whether these are professional experts in public administration problems or not.

4. Conclusion

An important function of regional informational environment is providing quality outputs on even more complicated or inaccurately formulated enquiries. Enquiry creators can be public administration personnel, staff of other organizations and especially citizens, i.e. enquiries are put with different accuracy, different knowledge in information technologies and different knowledge in public administration problems. Therefore, the solution is a preparation of a suitable informational environment, which would take given problems into account. Regional data warehouse should serve for analysing and answering strategic regional enquiries, so that this informational environment would serve for an effective area administration and also as a resource for citizens.

References:

[1] Ministerstvo informatiky CR. URL: <http:// www.mcr.cz>. (Ministry of informatics Czech republic)

[2] Šimonová S., Komárková J. aj. *Conception of the regional development of the microregion Pardubicko.(original in Czech)* Pardubice: Univerzita Pardubice, 2001. 98 s. AV-Upa-FES-ÚSII-08.01/02. ISBN 80-7194-373-8.

[3] Terano T., Asai K., Sugeno M.: *Fuzzy Systems Theory and its Application.* Academic Press Limited London 1992

[4] Humphries, M., Hawkins, M, W.,: *Data warehousing – Architecture and implementation (Czech translation)* Copmputer Press Praha 2002

[5] Fabián P., Capek J.: *Data Quality, e-Government and e-Commerce*. Sci. Pap. Of the Univ. Pardubice Ser. D, 7. p. 29-34, (2002).