Fuzzy approach for searching in CRM systems

BOGDAN WALEK, JIŘÍ BARTOŠ, CYRIL KLIMEŠ
Department of Informatics and Computers
University of Ostrava
30. dubna 22, 701 03 Ostrava
CZECH REPUBLIC
bogdan.walek@osu.cz, jiri.bartos@osu.cz, cyril.klimes@osu.cz http://www.osu.cz

Abstract: This paper is focused on design of methodology for searching relevant data in CRM systems based on vague input terms. The paper proposes methodology for creating an expert knowledge base and a tool for processing vague input terms and extracting relevant data and also presenting these data as evaluated search results and sorting them according to given significance and vague input search criteria. The proposed approach for searching in CRM systems under uncertainty is verified on searching relevant data from the selected CRM system.

Key-Words: fuzzy; uncertainty; customer; CRM; expert system; searching; search results.

1 Introduction
In present days the concept of Customer Relationship Management (CRM) is very popular and frequently used. Customer Relationship is a comprehensive approach for creating, managing and expanding customer relationships and it is the basis of marketing and sales branches [1], [2] [6]. CRM describes a model for managing interactions and relationships between a company (an organization) and a customer [5]. CRM systems are used to organize and automate business processes for marketing and sales activities, customer service and support, and other activities that help companies to increase sales and profit [7]. The main goals of CRM are:

• to care about current customers,
• to find and win new customers,
• to regain former clients,
• to reduce the cost of marketing,
• to improve efficiency of business processes in company,
• to track product sales,
• to track projects and their activities.

Currently there are many implementations of CRM systems that work with large amounts of data about projects, customers, products, sales, etc. During the search process in these data it is appropriate to find the required data quickly and accurately. The article describes one of the possible ways of finding relevant data in CRM system.

2 Problem Formulation
To search and analyze relevant data belonging to customer, product or project full-text searches, filters or data mining methods are usually offered by CRM systems. However, these methods are unsuitable for processing uncertain data that consist either in searching criteria or in data of the system itself [3]. Sales managers, marketing managers and top management on the one side often use vague terms to describe information that they want to find in the CRM system, and on the other side, the relevance of the data is fixed. Working with these vague terms is more difficult, but it is appropriate to process vague terms and to extract relevant data and afterwards present various searching scenarios and views on the resulting search outputs.

3 Problem Solution
In this paper a fuzzy approach for searching relevant data in CRM system based on vague input terms and expert knowledge base is proposed. Knowledge base consists of IF-THEN rules which are based on vague input terms.

Proposed approach consists of the following steps:

1. Definition of desired search results
2. Definition of criteria for searching
3. Creating the knowledge base
4. Loading the relevant data from database
5. Processing input vague terms and evaluation of search results
6. Showing evaluated search results
Proposed approach is visually represented in the following figure:

Fig. 1 Proposed fuzzy approach for searching in CRM systems

3.1 Definition of desired search results
In the first step, definition of desired search results which will be shown to user is needed. There are many search results that can be shown to manager, managers and other users who are working with CRM system:

- Most significant customers
- Most profitable products
- Best products
- Most profitable periods, etc.

3.2 Definition of criteria for searching
Subsequently, it is necessary to define criteria that will be used for searching. Based on this information, the creation of the knowledge base that contains IF-THEN rules for evaluating search results will be possible. It will also be possible to create and compose appropriate query for retrieving relevant data from database. Here are a few possible criteria for searching the most significant customers (criteria were defined in section 3.2):

- Number of employees – information about number of employees of the specific customer (company), loaded from database table customers
- Products quantity – number of products that were purchased by specific customer, loaded from table products with identifier of the specific customer
- Products price – total price of all purchased products by specific customer, loaded from table products with identifier of the specific customer, prices of founded records are summed together
- Frequency of purchase – number of orders realized by specific customer in specific period, loaded from table orders

3.3 Creating the knowledge base
After definition of desired search results and criteria for searching, the knowledge base can be created. Proposed knowledge base consists of IF-THEN rules which is composed from defined criteria and determine degree of relevance of search result. Each criterion represents one input linguistic variable of IF-THEN rule. Output linguistic variable of IF-THEN rule is represented by degree of relevance of search result. Degree of relevance of search result is described by appropriate linguistic value which represents degree of membership of an item in set using membership function:

\[ A: X \rightarrow [0,1] \]

3.4 Loading the relevant data from database
Next, based on criteria defined in section 3.2, the relevant data will be loaded from database. The type of database that is used is relational database [8], [9]. For each desired search result is necessary to load partial data which will be used for finding the most appropriate IF-THEN rule from knowledge base. Afterwards, founded IF-THEN rule will determine degree of relevance of search result. Loading the relevant data from database will be described on searching the most significant customers (criteria were defined in section 3.2):

3.5 Processing the input vague terms and evaluation of the search results
In this step input vague terms (declared in section 3.2) will be processed and then search results will be evaluated by expert system (that contains knowledge base created in section 3.3 and uses appropriate IF-THEN rule. Here are few examples of input terms (desired search results) vaguely defined by user:

- Most significant customers
• Significant customers
• Very profitable products
• Good products
• Worst products

3.6 Showing evaluated search results
Finally, desired and evaluated search results are shown to user. Search results are sorted and ordered according to evaluation generated by expert system. For example in case of searching the most significant customers, the most significant customers will be shown at the top of the list with results and the least significant customers will be shown at the bottom of the list with results.

4 Results
Proposed approach will be verified on searching most significant customers which are stored in CRM system. Verified CRM system stores information about customers, products, purchases, employees, etc.

4.1 Definition of desired search results
In first step user defines that he wants to show the most significant customers that will be sorted by significance. Search results are declared as “find the most significant customers”.

4.2 Definition of criteria for searching
In next step criteria for searching the most significant customers were defined:
• Number of employees
• Products quantity
• Products price
• Frequency of purchase

4.3 Creating the knowledge base
Then the knowledge base is proposed and constructed. Input linguistic variables are automatically loaded from the previous step. Input linguistic variables:
1. Number of employees (EMPLOYEES) – example of values: few, medium, many.
2. Products quantity (PRODUCTS_QUANTITY) – example of values: small, medium, big.
3. Products price (PRODUCTS_PRICE) – example of values: low, medium, high.

4. Frequency of purchase
(FREQUENCY_OF_PURCHASE) – example of values: nearly never, rarely, often, very often.

Output linguistic variable:
1. Significance of customer
(CUSTOMER_SIGNIFICANCE) – example of values: low, medium, high, very high.

Vague descriptions of linguistic variables are represented by fuzzy sets. One of them is shown in the following picture (see Figure 2) – x-axis represents range of the interval, in which fuzzy sets interpreting the input variable are defined (lower and upper bound); y-axis defines membership degree of the input variable in fuzzy set (evaluation). Figure 2 represented the input linguistic variable FREQUENCY OF PURCHASE, which has 4 possible input attributes, attribute nearly_never is highlighted (left supreme=1, left kernel=5, right kernel=8 and right supreme=12).

![Fig. 2 Definition of fuzzy set for linguistic variable FREQUENCY_OF_PURCHASE](image.png)

Based on input and output linguistic variables the IF-THEN rules are created. Here are few examples of IF-THEN rules created by expert:

IF (EMPLOYEES ARE MANY) AND (PRODUCTS_QUANTITY IS BIG) AND (PRODUCTS_PRICE IS HIGH)
AND (FREQUENCY_OF_PURCHASE IS HIGH)
THEN
CUSTOMER_SIGNIFICANCE IS VERY HIGH

IF(EMPLOYEES ARE MEDIUM)
AND (PRODUCTS_QUANTITY IS SMALL)
AND (PRODUCTS_PRICE IS VERY LOW)
AND (FREQUENCY_OF_PURCHASE IS VERY LOW)
THEN
CUSTOMER_SIGNIFICANCE IS LOW

IF-THEN rules are designed and created in LFLC tool, which is described here [4]. Visualization of IF-THEN rules created in LFLC software is shown in the Figure 3:

![Figure 3](image-url)

**Fig. 3** Definition of IF-THEN rules

4.4 Loading the relevant data from database

Next, based on criteria defined in section 4.2, the relevant data will be loaded from a database. The database consists of database tables for storing information about customers and their employees, products, orders, etc.

Part of database schema is shown in the Figure 4:

![Figure 4](image-url)

**Fig. 4** Part of database schema of CRM system

Based on criteria, the SQL queries for each criterion are created. Examples of SQL queries shows loading information about first customer in database table customers (CustomerId = 1).

1. Number of employees
SELECT employees FROM customers
WHERE id = 1;

2. Products quantity
SELECT sum(quantity) FROM products p, orders o, customers c WHERE o.product_id = p.id AND o.customer_id = c.id AND c.id = 1;

3. Products price
SELECT sum(price * quantity) FROM products p, orders o, customers c WHERE o.product_id = p.id AND o.customer_id = c.id AND c.id = 1;

4. Frequency of purchase – selected period is a month – May 2012
SELECT count(product_id) FROM products p, orders o, customers c
WHERE o.product_id = p.id AND o.customer_id = c.id AND o.date >= '2012-05-01' AND o.date <= '2012-05-31' AND c.id = 1;

Results of SQL queries for each customer (company) in database are shown in the Table 1, here are explanatory notes: CUST - customer, EMP -number of employees, PRQ – products quantity, PRP – products price, FoP – frequency of purchase:

**Table 1** Information about customers loaded from database

<table>
<thead>
<tr>
<th>CUST</th>
<th>EMP</th>
<th>PRQ</th>
<th>PRP</th>
<th>FoP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACompany</td>
<td>50</td>
<td>250000</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>ABCCompany</td>
<td>70</td>
<td>221000</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>CCCCompany</td>
<td>20</td>
<td>271000</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>ECompany</td>
<td>100</td>
<td>3775</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>DECompany</td>
<td>150</td>
<td>3971</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>FACompany</td>
<td>200</td>
<td>4346</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>ESCOMPANY</td>
<td>300</td>
<td>2806</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>BOCompany</td>
<td>650</td>
<td>49891</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>HECompany</td>
<td>850</td>
<td>49514</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>JICOMPANY</td>
<td>1000</td>
<td>49190</td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>
4.5 Processing the input vague terms and evaluation of the search results

Now, the results of SQL queries for each company are sent as input data to expert system created in LFLC software. Then, appropriate IF-THEN rules are used for evaluation of the degree of customer’s significance.

Evaluation of degree of significance of customers is shown in the Table 2, evaluation is a value from interval [0,1] transformed to percentiles:

<table>
<thead>
<tr>
<th>CUST</th>
<th>EMP</th>
<th>PRQ</th>
<th>PRP</th>
<th>FoP</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACom pany</td>
<td>few</td>
<td>low</td>
<td>small</td>
<td>nearly never</td>
<td>2.3%</td>
</tr>
<tr>
<td>ABC o mpany</td>
<td>few</td>
<td>low</td>
<td>small</td>
<td>nearly never</td>
<td>2.7%</td>
</tr>
<tr>
<td>CCC o mpany</td>
<td>few</td>
<td>low</td>
<td>small</td>
<td>nearly never</td>
<td>4.6%</td>
</tr>
<tr>
<td>ECom pany</td>
<td>few</td>
<td>small</td>
<td>medium</td>
<td>rarely</td>
<td>29 %</td>
</tr>
<tr>
<td>DECo mpany</td>
<td>few</td>
<td>small</td>
<td>medium</td>
<td>rarely</td>
<td>28.7 %</td>
</tr>
<tr>
<td>FAC o mpany</td>
<td>medium</td>
<td>small</td>
<td>medium</td>
<td>rarely</td>
<td>28.3 %</td>
</tr>
<tr>
<td>ESC o mpany</td>
<td>medium</td>
<td>small</td>
<td>medium</td>
<td>rarely</td>
<td>30.3 %</td>
</tr>
<tr>
<td>BOCo mpany</td>
<td>many</td>
<td>big</td>
<td>high</td>
<td>very often</td>
<td>89.7 %</td>
</tr>
<tr>
<td>HEC o mpany</td>
<td>many</td>
<td>big</td>
<td>high</td>
<td>very often</td>
<td>90.3 %</td>
</tr>
<tr>
<td>JIC o mpany</td>
<td>many</td>
<td>big</td>
<td>high</td>
<td>very often</td>
<td>90.7 %</td>
</tr>
</tbody>
</table>

5 Conclusion

The result of this paper is a comprehensive model representing the search methodology using vague input terms. The methodology that uses fuzzy approach for searching relevant data in CRM system was proposed together with steps to construct the knowledge base that is used to evaluate the significance of the results and the fuzzy evaluation process itself.

The proposed model was verified and some of the results were presented.

Increased involvement of this tool in the other problem domains is also a topic of further research direction.

6 Acknowledgment

Presented topic is also a part of internal grant SGS10/PřF/2012, called Fuzzy modeling tools for analysis and design of information systems, at Department of Informatics and Computers, University of Ostrava.

References:


