







# Geographical Information System Based on Artificial Intelligence Techniques

Nayi Sánchez Fleitas<sup>1</sup> , Raúl Comas Rodríguez<sup>2</sup> ,  
María Matilde García Lorenzo<sup>1</sup> , and Frankz Carrera<sup>2</sup> 

<sup>1</sup> Universidad Central “Marta Abreu” de Las Villas,  
50100 Santa Clara, Villa Clara, Cuba  
nayi78@gmail.com, mmgarcia@uclv.edu.ec

<sup>2</sup> Universidad Regional Autónoma de los Andes,  
Ambato 180101, Tungurahua, Ecuador  
raulcomasrodriguez@gmail.com, frankzcarrera@gmail.com

**Abstract.** The Electrical Union in Cuba develops the Business Management System of the Electrical Union (SIGE) that focuses on the automation of electrical processes. The geographic information systems (SIGOBE) developed don't meet the specific requirements for their generalization due to their limited updating facilities and the small spectrum they cover. The general objective of the research is: to develop the geographic information system of the transmission and distribution processes in the Electric Union, with the use of artificial intelligence techniques, on a deep conceptual scheme of the domain, that responds to the requests of consultation of users as support for decision making. A case-based system on type problem solved was designed, using as an initial case database, the 265 static queries registered in SIGERE. The queries are described by eight data-type predictive traits and three objective traits. The similarity between two cases was determined by the weighted sum of the distance of their traits and the calculation of the distance between traits was done according to its nature. An intelligent real-time queries system was implemented for the SIGOBE, achieving the generation of automatic queries that allow the system to respond to any type of queries in real-time. The experimental study shows the feasibility of the proposal.

**Keywords:** Geographical information system · Artificial intelligence  
Case-based system

## 1 Introduction

Energy is one of the supports of development in production processes, social progress and technological progress [1, 2]. In this sense, technology has allowed to increase the decisions to make in the energetic production chain. In this environment, the stakeholders need management systems that serve as support for the decision making process, to ensure a more efficient sector [3].

The Electrical Union (UNE, Spanish acronym) in Cuba develops the Business Management System of the Electrical Union (SIGE) for the automation of electrical

processes [4]. SIGE is composed of two main subsystems: the Integral System of Network Management (SIGERE) and the Integral Management System of the Electrical Industry Construction Enterprise (SIGECIE).

The functions of SIGERE and SIGECIE are to collect technical, economic and management data to convert them into information. The data collected facilitate and improve the efficiency in the analysis, planning, operation, and control of the distribution and transmission electricity networks. Both systems establish the databases of a Geographic Information System (GIS) of the SIGE.

SIGERE and SIGECIE are considered complex systems because they have 36 modules and a database of: 716 tables, 1303 stored procedures and 74 functions. In addition, other functionalities are in development phase. An average action in the system involves approximately nine tables with different attributes. To carry out a query on a specific topic requires knowledge of the database organization. Despite the number of stored queries, they still do not cover the needs of the customer due to the operational dynamics of the national electro-energy system.

To solve this problem, an analysis of the literature is carried out and a group of experts on the subject is gathered. As a result of the previous analysis is decided:

- Conduct a national tour of all the electric companies and groups of the ECIE for the study of possible solutions.
- Search for information on the subject in the literature.

The geographic information systems developed for the electricity companies in the country do not meet the specific requirements for their generalization due to: their limited updating facilities and the small spectrum they cover. The main limitations found in the analyzed solutions are: fixed queries that are limited, the use of proprietary software and the short validity of the data that are inserted into the map. An investigation is made of existing GIS in the world and the current trends of its development.

Information technologies have evolved towards the construction and implementation of intelligent systems [5–7]. In the 1970s, computerized systems that use knowledge about a domain to arrive at a solution to a problem; and the knowledge and the solution method are separated, are called Knowledge Based Systems (KBS) or Expert Systems (ES). This solution is essentially the same as that obtained by a person experienced in the domain if faced with the same problem.

The general objective of the research is: to develop the geographic information system of the transmission and distribution processes in the Electric Union, with the use of artificial intelligence techniques, on a deep conceptual scheme of the domain that responds to the requests of consultation of users as support for decision making.

## 2 Analysis of the Methodological Basis

Artificial Intelligence (AI) is the branch of computer science that attempts to reproduce the processes of human intelligence through the use of computers [8]. Within AI, the Expert Systems (ES) or Knowledge Based Systems (KBS) emerges as the field in charge of the study of: the knowledge acquisition, its representation and the generation of inferences about that knowledge. There are different variants to build the KBS based

on the representation of knowledge and the method of inference to implement. Among the systems are: Rule Based Systems [9, 10], Frames Based Systems [11], Probability Based Systems [12], Expert Networks [13, 14], and Case-Based Systems or Case-Based Reasoning Systems (CBR) [15–18].

In this sense, CBRs appear as a palliative to the process of knowledge engineering and are based on the premise that: similar problems will have similar solutions [19, 20]. With this principle as the basis, the solution to a problem is retrieved from a memory of solved examples. For each case, the most similar previous experiences that allow finding the new solutions are taken into account [21, 22].

The CBRs need a collection of experiences, called cases, stored in a case database, where each case is usually composed of the description of the problem and the solution applied [23]. Case-based reasoning contributes to progressive learning, so that the domain does not need to be fully represented [24]. The CBRs have three main components: the user interface, the knowledge base and the inference engine [25, 26].

## 2.1 Case Database

A case contains useful information in a specific context, the problem is to identify the attributes that characterize the context and to detect when two contexts are similar. Kolodner defines that “a case is a contextualized knowledge that represents an experience” and it is described by the values that are assigned predictive and objective traits [27].

To provide the system with a conceptual basis the traits can be organized through ontology. The fundamental role of ontology is to structure and retrieve knowledge, to promote its exchange and communication [28–30]. In addition, relying solely on CBR for distributed and complex applications, can lead to systems being ineffective in knowledge acquisition and indexing [31]. According to Bouhana, et al. [32] use of ontologies in case-based reasoning gives the following benefits:

- It is an easy-to-use tool for case representation.
- Queries are defined using daily terminology.
- It facilitates the assessment of similarity.
- It increases system performance.

A lightweight ontology is provided to the system to give a conceptual basis. In the conceptualization we have the concepts of taxonomy and relations (objects properties). The remaining components of the ontology model (data properties, instances and axioms) aren’t developed because the information is already in the database that feeds the system. The ontology was carried out using the Methontology methodology and the Protégé tool [33].

With the ontology the system gives a conceptual basis. Nevertheless, a weakness of the systems proposed is in the dissatisfaction with the queries carried out. If a static inquiry is developed for each problem that arises, the database begins to store a group of scarcely-used queries. In order to solve the problem, the system must be able to generate intelligent queries in real-time, in which the knowledge obtained from previous ones is used.

The success of the Case Based Systems (CBS) is achieved in the knowledge acquisition module because: the basic elements of the problem or domain to be solved

are described, and the way to facilitate the representation of the knowledge of the expert is looked [34]. To form the Case Based (BC), systems and subsystems that generate data or information in the knowledge acquisition process must be analyzed.

This domain contains alphanumeric and geographic information. The geographic data correspond to the cartography that can be downloaded from the internet or buy from an enterprise. With the data it obtain the spatial position of the object. The elements of the monolineal are the basic elements of the electrical network that are represented in correspondence with their characteristics in the vectorial scheme, which occupy a position in the geographical space (Table 1).

**Table 1.** Representation of the network in vectorial scheme.

Points (support points)	Lines (circuit)	Polygons (substation)
Posts	Transmission circuit	Distribution substation
Transformer banks	Subtransmission circuit	Transmission substation
Capacitor banks	Primary circuit	
Generator groups	Secondary circuit	
Disconnectors	Lighting circuit	
Structures		
Lamps		

The data provided by the SIGERE and SIGECIE modules are stored in the database. The database of the SIGERE has 716 tables, 1303 stored procedures and 74 functions; therefore, any query can have a high degree of complexity. For example, a query action, to the transformer module, involves an average of nine tables and approximately 140 attributes. SIGERE and SIGECIE differ in the voltage levels they cover.

A simple query of the SIGOBE (E1), becomes three queries to the system (C1, C2, C3). The steps to perform the development of a query are: 1- (E1) Query entry; 2- (C1) Request of the user; 3- (C2) Query of the database; 4- (C3) Query to the cartographical data; 5- Show C3; 6- Build legend; 7- Locate selection and show the legend.

In the case of a query that has two or more conditions  $x_1, x_2, \dots, x_n$ , then  $Q(x_1, x_2, \dots, x_n)$  provides a solution set of S. If the query is decomposed into n simple queries  $Q(x_n)$ , which provide n solution sets  $S_1, S_2, \dots, S_n$ ; then, the solution set corresponds to the union, intersection or the Cartesian product of these sets.

$$Q(x_1, x_2, \dots, x_n) = S \rightarrow Q(x_1) \text{ op } Q(x_2) \text{ op} \dots \text{ op } Q(x_n) = S \tag{1}$$

Where

$Q(x_1, x_2, \dots, x_n)$ : query.

$x_1, x_2, \dots, x_n$ : terms.

op: operator ( $\cap/\cup/X$ ) of the sets resulting from the queries.

S: solution set.

For the research, it's necessary to define a group of experts that provide data and knowledge. In the creation of the group, we work with the methodology proposed by [35], the number of experts needed is calculated using the probabilistic method of the binomial distribution coefficient, an initial list of possible experts is determined and a questionnaire is applied to determine the competence coefficient in correspondence with: years of experience, levels of knowledge and source of acquisition.

To determine the structure of the case, the group of experts brainstorms and determines that each case is composed of eleven fundamental traits: eight predictive traits and three objectives. Figure 1 shows the structure of a case. In the Table 2 identifies the universe of discourse of the predictive and objective traits.

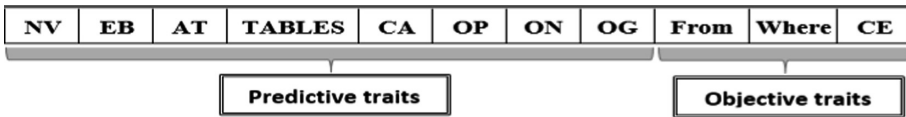


Fig. 1. Cases structure.

Table 2. Universe of discourse of predictive and objective traits

Trait	Possible values	Type
<b>Predictive traits</b>		
NV	Secondary, primary, subtransmission, transmission	Symbolic and single-valued
EB	Posts, transformer banks, capacitor banks, generator groups, disconnectors, structures, lamps, transmission circuit, subtransmission circuit, primary circuit, secondary circuit, lighting circuit, distribution substation, transmission substation	Symbolic and single-valued
AT	Attributes to be returned by the inquiry (code, voltage, name, etc.)	Set
Tables	Tables of the SIGERE involved in the inquiry (accessories, actions, connection, interruptions, line, currentSupplyPrimary, etc....).	Set
CA	Element to compare (attribute being compared)	Symbolic and single-valued
ON	Operator ( $\cup$ , $\cap$ , $\leq$ , $\geq$ , $=$ , like, etc.)	Symbolic and single-valued
OP	Ontology (descriptive logic) $T \cap TP \cap TMon-SSecon$	Ontology
OG	Spatial constraint (descriptive logic)	Ontology
<b>Objective traits</b>		
From	Returns the From of the inquiry to the SIGERE	String
Where	Returns the Where of the inquiry to the SIGERE	String
CE	Returns the GIS inquiry	String

The ON and OG ontological traits are represented by descriptive logic. A possible value of the ON trait would be:  $T \cap TPot \cap TMonophasic \cap SSecondary$ . This range expresses that the element is a monophasic primary transformer without secondary output.

OG works similarly, but their relationship is spatial, an example that refers to the location of an element would be:  $P \cap Prov \cap Muncp$ . This example expresses that an element belongs to the country (P), to a province (Prov) and to a municipality (Muncp).

As the ontology becomes deeper the element to be consulted fulfills requirements that can be flexible to the user. For one element to be similar to another it doesn't need to be at the same level in the ontology, but it must have gone through the same branch. The degree of importance of each level decreases as one goes deeper in the tree.

In the present investigation for the organization of the CB a hierarchical structure is used because: it favors the system, the access process and the recovery of the most similar cases. To do this, an analysis of the traits is made that allows to discriminate more options in each case.

Sixteen possible structures were tested (Table 3), for which the relationship between: the quality of the generator (percentage of well generated cases) and the complexity of the recuperator (percentage of recovered cases) is established. It was decided that NV was the root of the tree because of the importance of the trait. Additionally, since EB only belongs to one voltage level it does not make sense to put the NV below EB; the number of cases recovered does not change and the three structures contain the EB above the OP. Therefore, the best variant is NV - EB - OP (11), where the highest classification percentage is reached with the lowest percentage of recovered cases.

**Table 3.** Possible structures were tested.

Heading level	1 <sup>st</sup> level	2 <sup>nd</sup> levels	3 <sup>rd</sup> levels
Without structure (1)	NV (2)	NV - EB (5)	NV - EB - OP (11)
		NV - OP (6)	NV - OP - EB (12)
	EB (3)	EB - NV (7)	EB - NV - OP (13)
		EB - OP (8)	EB - OP - NV (14)
	OP (4)	OP - NV (9)	OP - NV - EB (15)
		OP - EB (10)	OP - EB - NV (16)

Figure 2 shows the hierarchical structure designed, where: NV is the root node; in the second lever EP and in the third level OP, for being discriminative elements. In the leaf nodes there is a subset of cases that represent the examples where the value of NV, EB and OP coincide, which reduces the search of cases in the database, by concentrating on a subset of cases that responds to the values given to the NV, EB and OP. In each case subbase there are 20 cases on average, which increases the speed of the recovery process.

In the present research there are three types of data: symbolic, set y ontologic; for which different distance measurements are used.

- NV, EB, CA and OP traits are symbolic and single-valued type, the distance used is Boolean [36].
- AT and Tables traits are of joint type and use Jaccard distance [37].
- The ON and OG traits represent the general and spatial ontologies and the distance used is Jaro-Winkler [38].

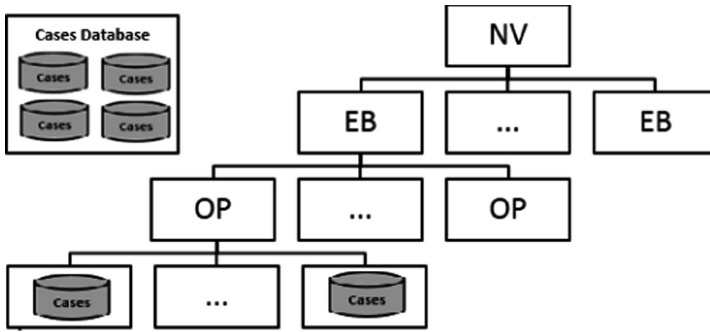


Fig. 2. Cases database structure for the UNE.

The recovery uses local comparison criteria or a local distance that determine the similarity or closeness between values of the same trait and a measure of global dissimilarity, which combines the results of the local criteria of all the predictive traits of the cases to be compared with the new problem.

The global dissimilarity measure used is the one proposed in [39]:

$$DisimGlobal(X, Y) = \left[ \sum_{i=0}^n w_i * d_i(x_j, y_j) \right] / n \tag{2}$$

The weights give a level of importance to each trait, the greater the weight the more important the trait. For the calculation of weights we work by means of expert criteria and the AHP method is used [40].

The objective is to recover the most similar k-cases, in a subset of the CB, which contain the cases with the same NV, EB and OP as the problem to be completed.

The correct dimension of the recovered cases was recalculated. Experiments were performed with odd k values from k = 1 to k = 11. As can be seen in Table 4, the results with k = 1 and k = 3 are very similar, but as the value of k increases, the percentage of correctly generated solutions decreases. It must be taken into account that the increase in the value of the k complicates the selection process of the initial solution.

Table 4. Possible K.

K	Correct generation	K	Correct generation
1	95.77%	7	89.41%
3	95.23%	9	83.49%
5	92.59%	11	78.83%

The recovered cases are combined together with the new case through the reuse of information. The solution proposal is made through similarity mechanisms that define the proximity or not of the recovered case, with the new one (Fig. 3).

In: P; problem to solve and  $BC' = \{bc_1, bc_2, \dots, bc_n\}$ ; sub-set of cases obtain of the sub-base that corresponds to the P case.  
 Out: Sets of K cases more similar.  
 $\forall bc_j \in BC'$  is calculated the  $SimGlobal(bc_j, P)$ . Where:  
 $SimGlobal(bc_j, P) = \sum_{i=0}^8 w_i * d_i(bc_{ji}, P_i) / n$

$$d_i(bc_{ji}, P_i) = \begin{cases} 0 & \text{Si } bc_{ji} == P_i & i = 1,2,5,6 \\ 1 & \text{Si } bc_{ji} \neq P_i & i = 1,2,5,6 \\ \frac{|bc_{ji} \cap P_i|}{|bc_{ji} \cup P_i|} & i = 3,4 \\ \frac{1}{3} \left( \frac{c}{|bc_{ji}|} + \frac{c}{|P_i|} + \frac{c - t/2}{c} \right) & i = 7,8 \end{cases}$$

The most similar K cases are returned with a  $K=3$ .

**Fig. 3.** Recuperator algorithm of k most similar cases.

Based on the transformational analogy, the Algorithm to propose an initial solution is developed. The new case is evaluated and adapted to the conditions on the recovered cases. The pre-set consultations are not necessarily identical to those stored in previous cases. To develop an initial solution, all recovered cases are considered and a combination of the recovered solutions is taken as a starting point.

## 2.2 Adaptation

The input of the adaptation module is an initial solution of the three objective traits. This module allows to reuse and adapt based on transformational analogy, which implies structural changes in the solution. Transformational adaptation is guided by common sense where the rules were defined and used in the adaptation process. This process is considered a T-space, where the known solution (KS) is going to be transformed with the use of T-operators (Table 5), until it becomes the solution of a new problem.

**Table 5.** T-operators according to the trait objective

From trait	Where trait	CE trait
Coupling insert	Insertion of restriction	Insertion of restriction
Coupling removal	Removal of restriction	Removal of restriction
Replacing coupling	Substitution in restriction	Substitution in restriction



Each T-operator is defined by a set of rules that perform the operation indicated. These rules perform a chain work that allows inserting, eliminating or replacing part of the solution to adapt it to the needs of the current problem. This work satisfies the restrictions imposed by the experts in domain ontologies and the natural requirements of objective traits. The adaptation module has three stages that are described below.

In stage 1 the review of the three objective trait has an algorithm of 25 rules that allow to check which traits are absent, which are valid and which ones need to adapt.

In stage 2, the set of rules to be applied is chosen according to the adaptation requirements of the previous stage in the following way:

- If there are no requirements, return the initial solution without adapting.
- If there is a requirement for the FROM trait, the set of rules is applied to adapt the FROM trait.
- If there is a requirement for the WHERE trait, the rule set is applied to adapt the WHERE trait.
- If there is a requirement for the CE trait, the set of rules is applied to adapt the CE trait.

In stage 3, a total of 65 adaptation rules are applied. The adaptation rules are divided into subsets by methods 5, 6, 7 and 8.

The methods 5 and 6 contain the set of 24 rules that allow you to adapt the FROM trait. This requirement can be given by:

- The initial result of the trait is absent since the FROM is an empty string or because  $\forall Xi \in (Tablas) | Xi \notin From$ .
- The solution does not contain the correct base table. The rules are responsible for finding the correct base table and replacing it in the initial solution. The replacement of the base table in the solution can lead to the previous one being related to it and replacing it would be a coupling of the new base table with it. The application of rule 7 eliminates this type of coupling.
- $\exists Xi \in (Tablas) | Xi \notin From$ . In this case there would be no tables in the result, so the solution would be incomplete and it is necessary to add them coupled to the base table.
- $\exists Xi \in (From) | Xi \notin Tablas$ . In this case, tables would be left over in the result and it is necessary to eliminate the coupling to the base table.
- After the base element is selected, the rest of the missing tables is added.

The method 7 contains the set of 27 rules that allow adapting the Where trait. These rules will be executed when the revision step establishes that the Where trait needs to be adapted. This requirement can be given by:

- The initial result of the Trait is absent because it is an empty string and the case has restrictions.
- The solution does not have the correct OP trait.
- The solution does not have the correct AC trait.
- The solution is not given according to the general ontology of the SIGERE system in the output, phase or type of correct installation.

The method 8 contains the set of 13 rules that allow the CE trait to be adapted. These rules will be executed when the revision step establishes that the CE trait needs to be adapted. This requirement can be given by the absence of any term of the spatial ontology in the query.

Once the adaptation and review of the expert, which confirms that the adaptation is correct, the case will be retained in the base of cases with the aim of enriching it with the solutions of new problems.

The retention is induced from the cases, so it will be necessary to redefine it periodically. The efficiency of the system is affected when the number of cases grows excessively, therefore, it is important to avoid including cases that do not contribute new information to the system.

To carry out the retention of cases, the following steps are followed:

- The degree of information provided by the case to the system is calculated. This degree of information is estimated by the number of T-operators applied between the set of T-operators in the T-space.
- It is considered feasible to retain the case whose degree of information provided is greater than  $\alpha$  (represents an information threshold).
- If the case is feasible, it is retained in the corresponding sub-base according to the value of the NV, EB and OP predictive traits, given the calculation of the degree of information provided by an objective trait.

The degree of information provided to the system by the value of an objective trait is calculated as the minimum number of T-operators applied.

### 3 Results

For the implementation of the CBR, the SICUNE module was developed. The SICUNE module has a national character and is applied in different areas of the electricity companies. This software can be considered as a support system for decision making because it fulfills the following characteristics:

- It's focused on the analysis of the operational data of Cuban electric companies.
- Performs dynamic, flexible and interactive reports, in correspondence with the needs of the decision-making areas in the Cuban electricity companies.
- It presents rapidity in the response time.
- It has availability of historical information, managed in the databases of SIGERE and SIGECIE for more than 15 years.
- It uses in the main areas of knowledge management in the Cuban electricity companies that need graphic modeling.
- Main areas of knowledge management are present in the electricity companies that need graphic modeling.

The incorporation of SICUNE into the SIGOBE increases the spectrum of search requests and provides a group of facilities such as: locating complaints from the population, a failed installation or with abnormal parameters, organizing the route of the cars, visualizing the voltages of the customers on the map, make a study of

equipment faults by zones, the optimization of the use and an optimal expansion of the networks, access to the information of a point of the distribution line, the study of electrical losses and certain scales allows to draw the sketch of the new projects with the necessary accuracy.

With the development of ontology, comprehensive access to the database and dynamic queries are achieved, which are fundamental in high demand stages of services due to the speed of the link between alphanumeric and graphic information. The processing of the technical language executed by the system increases the ease and convenience with which requests are made and allows the user ignore the structure of the database. In addition, concepts common in the domain that are not represented in the database are incorporated.

The efficiency is a component of productivity and is related to the use of inputs during the transformation process [41], so it can be evaluated taking into account the computational complexity that indicates the effort that must be made to apply an algorithm and how expensive it is.

In order to validate the investigation in terms of efficiency, the analysis of the computational complexity of the algorithm must be done. The worst case is the complexity of the reasoner without structure, that is to say, if the base of cases had a sequential structure. The computational complexity is calculated as:

$$N * ComplexityDisimGlobal(X, Y) \tag{3}$$

Where:

N: is the number of cases in the base of cases.

Complexity DisimGlobal (X, Y): it is the sum of the complexities of the local dissimilarity functions, in the case of a sum of temporal complexities, the greatest of all the complexities is taken as general complexity. The complexities of the local dissimilarity functions ( $n_1, n_2$ ) are:

- Jaccard’s temporal complexity ( $n_1$ ) (Used for set type attributes)  $O(n)$ , where n is equal to the number of elements of the attribute domain. Worst case  $n_1 = 90$ .
- Temporal Complexity of Jaro ( $n_2$ ) (Used for ontological traits)  $O(m)$ , where m is the length of the chain. (AMÓN, 2010). Worst case  $n_2 = 25$ .

So the complexity of the algorithm would always be an  $O(n)$ .

Complexity of the reasoner with structure, that is, the hierarchical structure:

$$\frac{N}{12} * ni \tag{4}$$

Where:

$N/12$ : will be the number of average cases recovered.

$ni$ : the greatest of the complexities of local dissimilarity functions.

So it would also be an  $O(n)$ .

A balance between precision and efficiency is achieved, because the model for data management, which allows smart consultations in the geographic information system for decision making in the transmission and distribution processes in the UNE

provides: a good performance of prediction, in a reasonable response time, with low computational complexity.

To test the SICUNE, three departments of an Electric Company, in a province, are selected that use information from different areas of the database and achieve greater coverage in the information contained. The work of these areas is operational and needs the functionality proposed in their daily work: Command post; Engineering department; Customer service.

A study of the exploitation of SIGOBE v1.0 determining their approximate use (Table 6).

**Table 6.** Queries to SIGOBE v1.0.

Queries	Use	Average of monthly consultations
Command post	See the monolineales Analyze the limits of freeways complaints	100
Engineering department	Analysis of the study area Transformer inquiries	25
Customer service area	Customer service State of complaints Status of projects and studies	50

In all three departments, the exploitation of the new version of SIGOBE begins, with the SICUNE module incorporated, for a period of one month for its validation. Table 7 shows the results by area.

**Table 7.** Results of SICUNE by area.

Queries	Total queries	Correctly classifies	Bad classifies	Retain	Effectiveness
Command post	230	221	9	23	96.08%
Engineering department	189	178	11	40	94.18%
Customer service area	147	140	7	18	95.23%

The engineering area was the least represented in the case based, because the SIGOBE, version 1.0, focused on the Command post, the attention to the complaints of the population and the investment area. However, 94.18% effectiveness is obtained. The percentage of cases solved incorrectly corresponds to the few cases on these areas that were counted at the beginning of its application.

## 4 Conclusions

- A case-based system on type problem solver was designed, using as an initial case database, the 265 static queries registered in SIGERE. The queries are described by eight data-type predictive traits and three objective traits. The case database responds to a three-level hierarchical organization, which favors the processes of access, recovery and learning of cases.
- Calculation of the distance between traits was done according to its nature. It was determined that the best results in the study case are: for the traits of nominal type, the Boolean distance; for traits of set type, the Jaccard distance and the ontologies were treated as strings using the Jaro Winkler distance.
- The case retention stage is in preliminary phase, since the current size of the case database does not presuppose reissues of cases, because it is still medium-sized.
- An intelligent real-time queries system is implemented for the UNE (SICUNE), achieving the generation of automatic queries that allow the system to respond to any type of queries in real time.
- The experimental study shows the feasibility of the proposal. The CBS obtains an effectiveness of 94.18%, where 539 searches of the total number of consultations were correctly classified. In the study 81 new cases were retained for a total of 301 in the case base. The percentage of cases solved incorrectly corresponds to the few cases on these areas that were counted at the beginning of its application.
- It's necessities to develop the axioms of the ontology for the data management of the processes of transmission and distribution of the UNE.
- Extend the SIG to the Generation process of the Electric Union including: Thermal Generation, Distributed Generation and Renewable Energy Sources Management.
- The study of possible applications of the model for data management in other branches of knowledge.

## References

1. Mónica Andrea Arango, A., Santiago Arroyave, O.: Análisis de combustibles fósiles en el mercado de generación de energía eléctrica en Colombia: un contraste entre modelos de volatilidad. *Revista de Métodos Cuantitativos para la Economía y la Empresa* (22), 190–215 (2016)
2. Castillo, Y., Castrillón Gutiérrez, M., Vanegas Chamorro, M., Valencia, G., Villicaña, E.: Rol de las Fuentes No Convencionales de Energía en el sector eléctrico colombiano. *Prospectiva* **13**(1), 39–51 (2015)
3. García, M.L.: Modelos de predicción de demanda eléctrica utilizando técnicas de inteligencia artificial. Aplicación al mercado eléctrico español. Universidad Miguel Hernández de Elche (2016)
4. Fernández Álvarez, R.: Informatización de la Gestión de las Redes Eléctricas, Tesis Doctoral, Facultad de Ingeniería Eléctrica. Universidad Central “Marta Abreu” de Las Villas, Santa Clara (2011)

5. Gondres Torné, I., Rodríguez León, N., Lajes Choy, S.E., del Castillo Serpa, A.: El aprendizaje bajo incertidumbre aplicado al mantenimiento de interruptores de potencia. *Ingeniería Energética* **XXXV**(2), 149–158 (2014)
6. Martínez, M., Santana, E., Beliz, N.: Análisis de los paradigmas de inteligencia artificial, para un modelo inteligente de gestión de la energía eléctrica. *Revista de Iniciación Científica* **3**(1), 77–84 (2017)
7. Rueda Calier, F., Rodríguez Suárez, A., Castellanos Granados, H.C.: Desarrollo y tendencias de la inteligencia artificial. *Revista Matices Tecnológicos*, 4 (2013)
8. Russell, S.J., Norvig, P.: *Artificial Intelligence. A Modern Approach*. Prentice Hall, Englewood Cliffs (1995)
9. Guerrero, O.E.G., Osorio, J.J.R., Gauta, T.L.A.: Implementación de una estrategia de control difuso para aumentar la producción de crudo en pozo petrolero. *Revista Colombiana de Tecnologías de Avanzada (RCTA)* **1**(27), 98–103 (2017)
10. Rey, H.C., Ricaurte, J.A.B., Pérez, A.G.: Sistema experto para la selección de personal desarrollador de software. *Ingenio Magno* **4**(1), 75–81 (2014)
11. Zhang, Y., Luo, X., Zhang, H., Sutherland, J.W.: A knowledge representation for unit manufacturing processes. *Int. J. Adv. Manuf. Technol.* **73**(5–8), 1011–1031 (2014)
12. Marling, C., Shubrook, J., Schwartz, F.: Case-based decision support for patients with type 1 diabetes on insulin pump therapy. In: Althoff, K.-D., Bergmann, R., Minor, M., Hanft, A. (eds.) *ECCBR 2008. LNCS (LNAI)*, vol. 5239, pp. 325–339. Springer, Heidelberg (2008). [https://doi.org/10.1007/978-3-540-85502-6\\_22](https://doi.org/10.1007/978-3-540-85502-6_22)
13. Amin, H.H., Deabes, W., Bouazza, K.: Clustering of user activities based on adaptive threshold spiking neural networks. In: *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 1–6. IEEE, Milan (2017)
14. Platon, R., Dehkordi, V.R., Martel, J.: Hourly prediction of a building's electricity consumption using case-based reasoning, artificial neural networks and principal component analysis. *Energy Build.* **92**, 10–18 (2015)
15. Cocea, M., Magoulas, G.: Design and evaluation of a case-based system for modelling exploratory learning behaviour of math generalisation. *IEEE Trans. Learn. Technol.* **14**(8), 1–14 (2017)
16. García Lorenzo, M.M., Bello Pérez, R.E.: A model and its different applications to case-based reasoning. *Knowl.-Based Syst.* **9**(7), 465–473 (1996)
17. González Benítez, N.: Sistema Experto basado en casos para el diagnóstico de la Fasciola hepática en el ganado bovino. *REDVET. Revista Electrónica de Veterinaria* **17**(12), 1–11 (2016)
18. Gutiérrez Martínez, I., Bello Pérez, R.E.: Making decision in case-based systems using probabilities and rough sets. *Knowl.-Based Syst.* **16**(4), 205–213 (2003)
19. Gillespie, K., Gupta, K.M., Drinkwater, M.: Case-based object placement planning. In: Lamontagne, L., Plaza, E. (eds.) *ICCBR 2014. LNCS (LNAI)*, vol. 8765, pp. 170–184. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-11209-1\\_13](https://doi.org/10.1007/978-3-319-11209-1_13)
20. Leung, K., Choy, K., Tam, M., Lam, C., Lee, C., Cheng, S.W.: A hybrid RFID case-based system for handling air cargo storage location assignment operations in distribution centers. In: *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 1859–1868. IEEE, Portland (2017)
21. de la Rosa Turbides, T.E.: Razonamiento basado en casos aplicado a la planificación heurística. Tesis Doctoral, Departamento de Informática, Universidad Carlos III, Madrid, España (2010)

22. Moreno Laverde, R., Joyanes Aguilar, L., Giraldo Marín, L.M., Duque Méndez, N.D., Tabares Morales, V.: Modelo para personalización de actividades educativas aprovechando la técnica de Razonamiento basado en Casos (RbC). *Campus Virtuales* **4**(1), 118–127 (2015)
23. Chazara, P., Negny, S., Montastruc, L.: Flexible knowledge representation and new similarity measure: application on case based reasoning for waste treatment. *Expert Syst. Appl.* **58**, 143–154 (2016)
24. Fitzgerald, T., McGreggor, K., Akgun, B., Thomaz, A., Goel, A.: Visual case retrieval for interpreting skill demonstrations. In: Hüllermeier, E., Minor, M. (eds.) *ICCBR 2015. LNCS (LNAI)*, vol. 9343, pp. 119–133. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-24586-7\\_9](https://doi.org/10.1007/978-3-319-24586-7_9)
25. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. *AI Commun.* **7**(1), 39–59 (1994)
26. Cordero Morales, D., Ruiz Constanten, Y., Torres Rubio, Y.: Sistema de Razonamiento Basado en Casos para la identificación de riesgos de software. *Revista Cubana de Ciencias Informáticas* **7**(2), 222–239 (2013)
27. Kolodner, J.: *Case-Based Reasoning*. Morgan Kaufmann Publishers, San Mateo (1993)
28. Myrgioti, E., Bassiliades, N., Miliou, A.: Bridging the HASM: an OWL ontology for modeling the information pathways in haptic interfaces software. *Expert Syst. Appl.* **40**(4), 1358–1371 (2013)
29. Valaski, J., Malucelli, A., Reinehr, S.: Ontologies application in organizational learning: a literature review. *Expert Syst. Appl.* **39**(8), 7555–7561 (2012)
30. Wimmer, H., Rada, R.: Good versus bad knowledge: ontology guided evolutionary algorithms. *Expert Syst. Appl.* **42**(21), 8039–8051 (2015)
31. Akmal, S., Shih, L.-H., Batres, R.: Ontology-based similarity for product information retrieval. *Comput. Ind.* **65**(1), 91–107 (2014)
32. Bouhana, A., Zidi, A., Fekih, A., Chabchoub, H., Abed, M.: An ontology-based CBR approach for personalized itinerary search systems for sustainable urban freight transport. *Expert Syst. Appl.* **42**(7), 3724–3741 (2015)
33. Sánchez Fleitas, N., Comas Rodríguez, R., García Lorenzo, M.M., Riverol Quesada, A.: Modelo de manejo de datos, con el uso de inteligencia artificial, para un sistema de información geográfica en el sector energético. *Enfoque UTE* **7**(3), 95–109 (2016)
34. Montoya Quintero, D.M.: Modelo para la extracción de conocimiento de un experto humano en un sistema basado en conocimientos usando razonamiento basado en casos. Tesis Doctoral, Facultad de Minas, Departamento de Ciencias de la Computación y de la decisión, Universidad Nacional de Colombia, Medellín, Colombia (2015)
35. Comas Rodríguez, R.: Integración de herramientas de control de gestión para el alineamiento estratégico en el sistema empresarial cubano. Aplicación en empresas de Sancti Spiritus. Doctor en Ciencias Técnicas, Departamento de Ingeniería Industrial, Universidad De Matanzas “Camilo Cienfuegos” (2013)
36. Althoff, K.-D., Auriol, E., Barletta, R., Manago, M.: *Review of Industrial Case-Based Reasoning Tools*. AI Intelligence, Oxford, USA (1995)
37. González Castellanos, M.: Extensión de algoritmos representativos del aprendizaje automático al trabajo con datos tipo conjunto. Tesis de Maestría, Departamento de Ciencias de la Computación, Universidad Central “Marta Abreu” de Las Villas, Santa Clara, Villa Clara, Cuba (2010)
38. Winkler, W.E.: Using the EM Algorithm for Weight Computation in the Fellegi-Sunter Model of Record Linkage. Bureau of the Census Statistical Research Division, Washington D.C., USA (2000)

39. Skjermo, J., Dah, E., Opland, R., Aamodt, A., Kofod-Petersen, A., Gustavsson, E., et al.: Case-based reasoning for alpine road operation support. In: 23rd ITS World Congress, Paper number EU-TP0079, Melbourne, Australia (2016)
40. Navarro Ferrer, D.G., Delgado, T., Martín, G., Jaimez, E.: An approach of a data-driven spatial decision support system to manage the effects of climate change on agriculture. In: 28th EnviroInfo 2014 Conference, BIS-Verlag, Oldenburg, Germany (2014)
41. González, C.A.S., Araque, D.P.: ¿ Entendemos realmente los conceptos entorno a la productividad?. *Revista San Gregorio* (17), 90–95 (2017)