Book: "**Expert System Software: Engineering, Advantages and Applications**"
Chapter: "**<u>A Practical Overview to Expert Systems in Telecommunication Networks, Medicine and Power Supplies</u>**"
Authors**: I.Monedero, A. Martín, J.M. Elena, J.I. Guerrero, F.Biscarri, C. León**

INTRODUCTION

Expert systems are being applied with success in multiple fields like engineering, medicine, geology, chemistry, etc., for the realization of diverse tasks (interpretation, prediction, diagnosis, design, planning, instruction, control, etc.). Some of these applications in these fields include:

- Inside engineering: The management and design of Telecommunication Networks.
- Inside medicine: The Disease Detection.
- Inside power supplies: Fraud Detection.

In this chapter, we will give an overview of the applications of Expert Systems to these fields. On the other hands, we will describe case studies to real problems solved in diverse projects carried out by the authors, explaining, from the experience of the authors, how to deal with such problems. The projects which will be described are the following ones:

- An expert system for the management of an ionospheric communication.
- An expert system for an efficient network management.
- An expert system for automatic routing of a HFC telecommunication networks.
- An expert system for the diagnosis of stomach disorders.
- An expert system for the detection of frauds in a power supply.

All these expert systems have been developed by the group of authors in the Electronic Technology Department of the University of Seville (Spain).

Inside Engineering: The Management and Design of Telecommunication Networks

As a result of the great evolution which has taken place in telecommunications due to the advances produced by new technologies, it is possible to access information rapidly, especially nowadays with the use of communication networks such as Internet. Telecommunication networks have evolved in time to satisfy the demands of different telecommunication services, which day by day require greater bandwidth and a better quality of service. Network technology has increased in complexity, generating the need for a better administration of the resources of these systems, which has favored the associated evolution of network management.

Expert systems may also be applied to the field of telecommunication networks, because many of the previously enumerated tasks are carried out in these networks. In addition, these tasks require specialized knowledge and, therefore, an expert system can help with the automation process.

Figure 1 schematically shows the different application domains of expert systems related to telecommunication networks [1][2]:

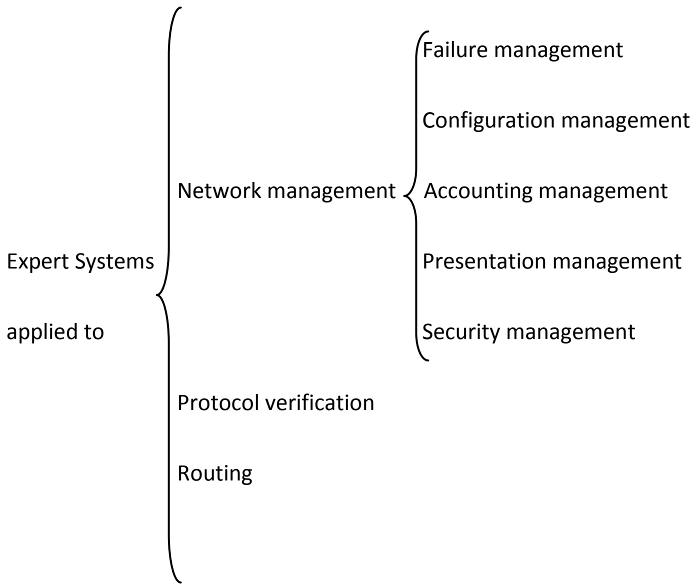

Figure 1: Expert systems in telecommunication networks

The efficiency of the tasks of operation and maintenance of current telecommunication networks largely depends on the degree of co-operation between the management system and the human operators. The management system plays a very important part, informing on possible anomalies and executing the commands of control.

The different tasks involved in this management and for which currently exist commercial expert systems, are:

- Failure management. This task refers to the group of functions related to network management which are necessary for the detection, diagnosis, recovery and correction of shortcomings of the elements that form a telecommunication network. Examples include Max & Opti-Max (an expert administrator of telephony line maintenance developed by NYNEX) [3], Trouble Locator (an intelligent system developed by Pacific Bell for locating failures at the physical level of a local telephony network) [4], ANSWER (Expert system developed by AT&T and in charge of supervising their 4ESS switches) and Scout (developed by AT&T for identifying persistent failures in the network) [5].

- Configuration management. This function helps the administrator to exercise control on the configuration of the components of the network. Changes in the configuration are carried out to eliminate congestions, to isolate shortcomings or to make changes which the users need. Examples of expert systems supporting configuration management include ECXpert (detection and reparation of failures, developed by Lucent Technologies) [5], ACE (an expert system developed by AT&T for helping in the detection and diagnosis of failures in cables) [6] and NEMESYS (also developed by AT&T for avoiding network congestions)[5].

- Accounting management. This function allows us to determine and locate positions and costs for the use of the resources of the network, e.g. APRI (also developed by AT&T for forecasting the probability of falling into new debts) [7].

- Presentation management. This task helps the network administrators to monitor and evaluate the benefits of the system. Example expert systems for presentation management are Net/Advisor & NetCommand (for monitoring the state of the network in real time) [5].

- Security management. This function helps the network administrators manage the services which provide protection in the access to the communication resources. Examples of expert systems related to security management include EXSYS (a set of expert systems developed by Pacific Bell which executes on an AT&T 3B2/600 computer for supervising the Loop Maintenance Operating System from a server) [5].

In addition, there are other expert systems such as NetHELP (which provides assistance for user trouble shooting) and ExSim (assistance in the process of routing a network) applied to other fields [5], enumerated in Figure 1.

There is an especially interesting yet less investigated field for the application of expert systems to the world of telecommunication networks: Systems applied to assist during the tasks of network design. Within the design process, the use of knowledge-based systems is of particular interest for the topological design of the network. Among these tasks are:

- The choice of the type of transmission: Transmission by physical support (cable, microstrip, wave-guided) and/or transmission by radio (electromagnetic waves).

- The layout of the network: On radio networks it would be necessary to establish the layout depending on the type of communication (bidirectional or omni-directional) and signal propagation and interference characteristics. On cable networks, this task comprises the process of wiring.

- Placement of the network elements: Establishment and localization of the different nodes and necessary elements in order to obtain the suitable signal level (splitters, amplifiers, etc).

Especially the latter two tasks require the work of experts who apply various design rules and their empiric knowledge. Because of the great number of limitations of the physical media, these tasks are especially complex in cable networks.

On the other hand, due to the very particular characteristics of the design rules for each network, as well as due to the high complexity of the application of the knowledge, the development and commercial use of expert systems for cable network design is still not very widely extended. Thus, the currently existing approaches mainly focus on research [8][9][10]. In some cases of paper [8], the design is only partial; others [9][10] are limited to the design of local networks.

Inside Medicine: Detection of Disease

The main advantage of expert systems in medicine [11][12] lies in making decisions in diseases with complex diagnosis [13][14]. These expert systems facilitate the work of doctors while producing an increase in productivity. Another benefit occurs when confirming a suspected diagnosis.

In order to realize an adequate diagnosis the doctor has a great amount of knowledge about the signs and symptoms of diseases. Thus, his/her expertise lies in the ability to relate the status of a patient with descriptions provided by the texts. The doctor determines what symptoms and signs are present and matches their meaning with the absentee ones and the disease with which he thinks that it is. The greater your ability to combine their knowledge with actual observations, the more accurate is his/her technique of diagnosis. The limiting factors are the ability to remember things organized, to correlate the observed cases with the pattern of existing data and apply this knowledge when the data are incomplete or not well adapted to the previous cases. The first two factors are the main advantage of computers; on the other hand, the latest factor is the strong point of human experts. If a computer system can replace the sharpness of the statistical analysis, its higher powers to organize data could enable it to emulate in some degree the doctor´s expertise. It is intended that the expert system is capable of approaching the diagnosis and help in management issues.

Expert systems have the capacity to learn, understand and be understood, to develop alternative hypotheses to resolve conflicts and to justify their conclusions. They have advantages such as handling a large amount of information, be useful in medical education because to develop them it is made thorough a study of the representation and reasoning of medical knowledge. In order to solve problems, expert systems are based on methods that derive their results by reasoning from knowledge about the specific topic previously entered into the system, in contrast with conventional systems that solve problems through algorithms or fixed repetitive processes, scheduled in advance and expect the same type of input data to process and generate the response. Expert systems are able to provide an explanation of the procedure followed to arrive at a solution, justifying the steps used in the process of deduction. Finally, expert systems can increase their knowledge base by adding more data. Among the drawbacks and disadvantages of expert systems we have to include the high cost of design and that its realization requires a highly specialized staff with sophisticated technology during a relative long time and long periods of testing to prove its reliability, especially when the system handles large volumes of information. Also an inappropriate use can have a negative impact on the relation between doctor and patient.

Since the early seventies diverse expert systems applied to medicine have been developed. For instance, the CASNET [15] is an expert system for diagnosis and management of glaucoma, the MYCIN [15] for diagnosis and treatment of infectious diseases, or the DIALOG-INTERNIST (today called Caduceus) to make differential diagnosis in the broad context of internal medicine. A more recent approach is ONCOCIN [16], which is focused on the topic of chemotherapy of patients with lymphoma.

On the other hand, many other applications with well defined roles and goals in different medical specialties have been developed. Each of these expert systems has a specific way of representing medical knowledge and of performing the logical reasoning to reach a conclusion and support it. The development of expert systems in medicine has received contributions from non-medical programs but which led to new ideas, among them the DENDRAL, META-DENDRAL and PROSPECTOR [15].

Inside Power Supplies: Fraud Detection

The application of expert systems to power utilities covers a wide spectrum of applications. They are applied in power energy production, transport, distribution and consumption. In some works [17][18] a series of categories very similar among themselves are proposed. These categories allow classifying functionalities of the expert systems in power systems applications:

- Alarm processing [19].
- Fault diagnosis or Technical losses diagnosis [20].
- Steady-state security and dynamic security assessment [21].
- Remedial controls [22].
- Restoration [23].
- Environments for operational aids [24].
- Substation monitoring and control [25].
- Maintenance scheduling [26].
- Expert System development methods and tools [27].

Smart Grids and remote management allow new research and application fields:

- Management of communication networks in power distribution networks [28][29].
- Management of electric vehicle or Vehicle-to-Grid (V2G) technology [30].

Normally, the current expert systems could be classified in more than one of these categories, because they aren't perfectly delimited and some expert systems have several functionalities.

In addition, customers' consumption is very important for power distribution companies. The non-technical losses are the main problem of the power distribution companies, because they cannot be forecasted as the technical losses. The techniques and technologies used on non-technical losses detection and classification are very similar to the ones used in other research fields, as for example: credit card fraud, intrusion detection, communication network fraud, etc.

The work [31] uses rough sets for fraud detection in electrical energy consumers. The technique used is based in the treatment of incomplete information about consumption. The successful of this one is of 20 %. On the other hand, the work [32] uses a Radial Basis Function (RBF) neural network. This method uses a matrix of 12 columns and n rows (according to the annual evolution of each of the variable on monthly periods) as input of the neural network.

In some case, as the work [33], the neural network is used to predict electrical energy consumption using the seasonal ARIMA model and MAPE (Media Absolute Percentage Error) index to compare the efficiency with other methods. But in this case, the demand forecasting is made over several customers. In the works [31] and [32] the detection is applied individually in each customer. In a same way as [33], the work [34] uses a Multi Layer Perceptron (MLP) which processes information about customer consumption get from ANOVA.

The work [35] combines a Support Vector Machine (SVM) and customers' consumption who have registered 25 months of consumption. In addition, the work [36] improves the results using genetic algorithms and the customers' consumption obtained through electronic-Customer Information Billing System (e-CIBS).

AN EXPERT SYSTEM FOR THE MANAGEMENT OF AN IONOSPHERIC COMMUNICATION: ICARO

Introduction

The use of the ionized layer that surrounds the earth, the ionosphere, as a medium for the accomplishment of long distance radio communications is increasing, as a consequence of the appearance of HF band radio sets joining high technical specifications with low cost and great availability. Added characteristics are high portability, due to its limited weight, and a great independence of other external communications systems as satellites or repeaters.

Contrariwise, the variable behavior of the ionospheric conditions and the increasing number of radio stations using the same portion of the radio electric spectrum makes frequently necessary to be an expert radio operator in order to establish the wished communication. This factor, that raises notably the cost of the radio communications system, makes especially attractive to adopt an expert system control strategy.

The ionosphere is the name applied to parts of the upper atmosphere of the earth containing free ions in adequate quantity to affect the propagation of electromagnetic waves [37].

The ionosphere is structured in several layers called D, E, F1 and F2, with different radio waves reflecting properties. Because of these differences, the receiving station may get the information by several different paths from the transmitting station, as can be appreciated in Figure 2.

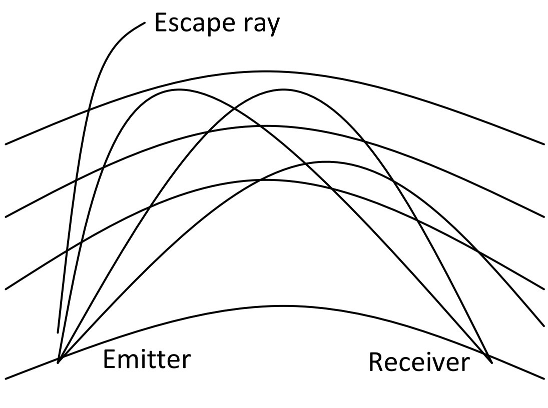

Figure 2: Different ionospheric paths

The most characteristic properties of the ionosphere as a communications channel are the long term variation, as the sun activity varies with a periodicity of approximately eleven years, the medium term variation, because the sun activity varies with the season of the year and finally the short term variation, that manifest the strong dependency between the hour of the day and the geographical coordinates of the stations intended to link.

It is also necessary to consider the influence of other different variables as the distance between radio stations, transmission power and the type of antennas, one of the link quality most determinant factors. The noise and previous occupation in the frequency that is wished to use is frequently the radio communication most limiting factors.

All these limiting factors imply that the radio operators must possess a great experience in the managing of their radio stations and a deep knowledge of all the elements implicated in the radio communication process.

The problem of the lack of expert operators is attempted to be solved through a new generation of radio communications equipment, capable of creating and maintaining automatic radio links, the HF ALE RADIOS (High Frequency Automatic Link Establishment Radios). Actually exist several federal standards (FS-1045A, FS-1046, FS-1047, FS-1048, FS-1049, FS-1050, FS-1051 and FS-1052) within the Federal Communication Standards Program for the U.S.A. National Communications Systems [38] [39].

These standards determine that an ALE radio communication set: automatically selects the best available channel based on link quality data stored in memory, establishes and confirm the links upon operator command, transfer data, do error checking, and relay messages.

Objective

The system (which was named ICARO) was conceived in order to facilitate the use of the long distance radio communications in the HF band to non-expert radio operators. It presents a user interface simple and effective and establishes a virtual link between the stations wished to be communicated.

As can be seen in Figure 3, it is realized with common elements generally present in any radio station: a transmitter receiver in HF, also denominated a transceiver, with its antenna and power supply and a personal computer. It is necessary to add a TNC (Terminal Node Controller) or controller for digital communications, in order to separate the error control function from the frequency managing function made in the PC computer. The TNC also include a multimode radio-modem and is an accessible element at an attainable cost. The HF transceiver, TNC, modem and keyboard are the required elements in a digital communications HF station. Just adding a personal computer running the ICARO program to a digital radio station determines its fully automation, so near to the ALE Radio standards as complete can be made the rules set of the expert system that govern it.

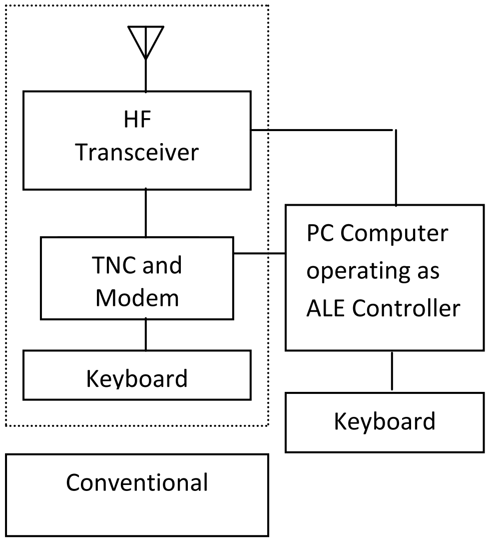

Figure 3: ICARO Radio System

The expert system uses different sources of data:

- Embedded data from an ionospheric forecast program. Usually these programs need to know [40] the location of transmitting and receiving stations, antenna types, time, day or month in the year, transmitter power, available frequency set, man-made noise level o ITU-R noise mode and SSN (Solar Spot Number) index, related with the solar activity. Given these inputs, the program habitually generates the HF path distance and great circle path bearings from site to site, the best usable frequency for each hour and the corresponding ray-path elevation angle, the received signal power, total noise power and signal/noise ratio.

- Information about propagation acquired by listening to automatically identified radio stations with known geographic coordinates. The system is in a continuous listening of a set of radio stations with frequency and geographic coordinates perfectly known. Every time a station is heard, the time, frequency and identification data is saved in a database register.

- Data from link quality analysis made in an automatic way when the stations are not required to do any other specific duty. If allowed by the radio operator, the system tries to connect to the previously listened radio stations that have radio transmission capacity. Provided the connection is made, a link quality analysis is realized, by the method of taking account of the time spent in transmitting successfully a determined amount of text.

The final objective of the system ICARO is to obtain automated digital communications following the basic behavior of ALE radio equipment but with a low cost, in contraposition to other systems in those which is intended the total integration of the calculation and radio communications capacity in a unique radio set. This last solution leads to expensive equipment of a very exclusive use, which makes difficult its employment in applications where the low level of utilization does not justify the high investment to accomplish.

Architecture

The forecast program proceeds from the ITU-R (International Telecommunications Union Radio Section). This program applies the UIT-R HF Propagation Prediction Method based on the ITU-R P533-5 recommendation [41]. It generates a file containing the everyday best frequency utilizable to communicate two geographic points over the earth surface. These frequencies belong to a frequency set previously determined for this type of digital communications.

The expert systems kernel is CLIPS (C Language Integrated Production System) from the NASA's Lyndon B. Johnson Space Center. This kernel provides a tool for handling a wide variety of knowledge with sup-port for three different programming paradigms: 1)Rule-based programming that allows knowledge to be presented as heuristics, or "rules of thumb", which specify a set of actions to be performed for a given situation. 2) Object oriented programming, allowing complex systems to be modeled as modular components, which can be easily reused to model other systems or to create new components. 3) Procedural programming possibilities, similar to the capabilities found in languages such as Java, Pascal, Ada or Lisp.

The rules that govern the expert system come from humane expert radio operators in the field of the HF ionospheric communications. These rules act on the transmitter-receiver to establish and maintain a digital link through an ionospheric channel.

A configuration program requests to the operator for the own station initial data and puts all the information in a database. This program will ask for the identification and geographic coordinates of the own station and identification call and geographic coordinates of the network stations and control stations, the type of antenna, etc. The configuration program will also asks for the identification calls and geographic coordinates of the network radio stations and control radio stations.

After the initial configuration process, the operator executes the ICARO principal program and the expert system creates the objects classes [BEST_ FREQUENCY] and [HEARD_STATION] and runs the forecasting program with successive entries, one for every network radio station. For each available frequency indicated by the forecasting program, an object of the [BEST_FREQUENCY] class will be instantiated with the attributes time and data of predicted connection, predicted best frequency and IRS Identification.

Later the expert system initiates the scanning process that tries to hear the network and control radio stations. The scanning process will be effectuated over the frequencies predicted by the propagation-forecast program, situated as an attribute in the objects belonging to the [BEST_FREQUENCY) class and establishing a short connection in order to perform a link quality analysis if the radio operator previously enabled this operation. Every time one of the expected radio stations is listened, an object of [HEARD_STATION] class will be instantiated, containing all the link attributes.

If authorized by the operator, the system will send periodically auto identification calls that, acting as a beacon, facilitates its localization by the other network stations. These calls will be emitted in the stand-by and scanning intervals of operation.

If the operator wishes to connect with a determinate radio station the system will consider itself as the ISS (Information Sending Station) and the other radio station as the IRS (Information Receiving Station). Also, it will ask the operator for the identification of the station desired to connect and will initiate the ionospheric forecasting program that will determinate the best frequencies, from the available frequency set, in order to establish the link.

The Expert System Rules

With all precedent information the expert system evaluates the possibility to establish the solicited connection, applying an adequate set of rules. Samples of these rules are:

1. Look for any object of the [HEARD_STATION] class with an attribute of station identity corresponding to the desired station and the attribute of time corresponding to the desired transmission hour. If it is found initiate the connection in the transmission frequency and at the hour indicated in the corresponding object attributes.

2. If there is not an object with the desired station identity and time attributes, or if it exist but the connection fails, look for any object of the [HEARD_STATION] class representing a control station with the attributes of been near the desired station and the wished transmission hour. If it is found initiate the connection in the transmission frequency and at the hour indicated in the corresponding object attributes.

3. After a previously determined period of time without getting connected, look for objects of the [BEST FREQUENCY] class with the desired time and identity attributes. If it is found send a connect-request call with the transmission frequency and at the hour indicated in the corresponding object attributes.

As can be observed the rules test the attribute values of the instantiated objects, triggering different actions in case of positive results in the comparison.

The rules elaboration was based in the sequence of operations that a human radio operator habitually follows. This knowledge comes from twenty written forms and ten personal interviews where expert radio operators tell his/her own experiences. After a trial period, a document with the behavior of the system was sent to the radio operators. Their appreciation and comments were used to evaluate, and modify, the rules set.

AN EXPERT SYSTEM FOR AN EFFICIENT NETWORK MANAGEMENT

Introduction

This system focuses on a framework and a language for formalizing knowledge management descriptions and combining them with existing Guidelines for the Definition of Managed Object definitions (GDMO). It was used an extension of GDMO standard with the following goals: facilitate the normalization and integration of the knowledge base of expert system into resources specifications, allow developers to specify the storage location and the update method of intelligent managed and provide a way to specify complex managed.

First it is analyzed the current management models, its evolution and the applications of the expert systems in the network management. It is carried out the normalization of the knowledge management necessary to administrate the resources that exist in the networks independently from the builder of the management resources. So, it was observed syntactically uniformed normalization of intelligence applied to the management. In the design it was used the new standard called Integrated Management Expert System that employs both managed model and AI reasoning techniques for the intelligent management of heterogeneous networks. This technique integrates the Expert System within the Management Information Base. The advantage is that a large problem can be broken down into smaller and manageable sub-problems/modules. For this goal a new property named RULE has been added, which gathers important aspects of the facts and the knowledge base of the embedded expert system. Secondary and in order to show the viability of own proposal, it was performed a practical demonstration in which the information and the management knowledge are unified in a unique specification, Figure 4.

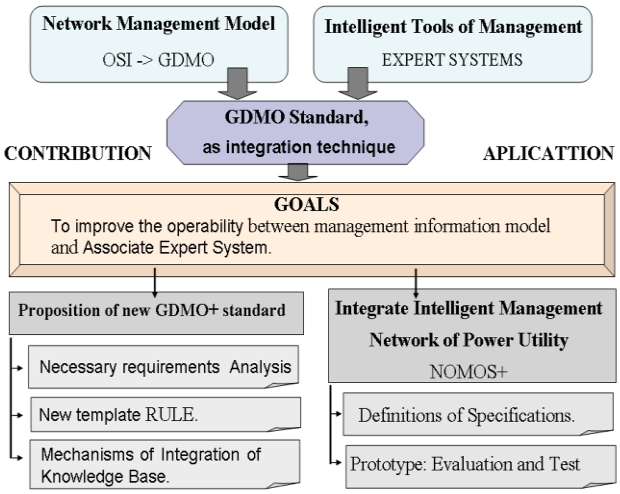

Figure 4: Synopsis of the research

<u>Management models</u>

A management model provides a common understanding between communicating processes in a heterogeneous environment, servers as the basis for a precise specification of network services and protocols and constitutes a vendor-independent concept. At the moment there are two main management models for computer communication:

- Internet: A structured and standardized approach to Internet management. This used the Simple Network Management protocol.
- The term Open System Interconnection (OSI) systems management actually refers to a collection of standards for network management that include a management service and protocol and the definition of a database and associated concepts: Common Management Information Protocol (CMIP), Systems Management Functions and a Management Information Model.

The fundamental function within OSI systems management is the exchange of management information between two entities: the managing system (the manager or requestor) and the managed system (the agent or responder) by means of a protocol CMIP [42]. The description of management information has two aspects. First, a Structure of Management Information

(SMI) defines the logical constitution of management information and how it is identified and described and second, the Management Information Base (MIB), which is specified using the SMI, defines the actual objects to be managed. The MIB is a conceptual repository of management information. It is an abstract view of all the objects in the network than can be managed [43].

Within the OSI management framework, the specification language GDMO has been established as a means to describe logical or physical resources [44]. GDMO has been standardized by ITU and is now widely used to specify interfaces between different components of the TMN [45]. This specification language allows network object designers and manager/agent implementers to communicate designs and build upon existing designs. Is the language used to define the structure and some of the relationships between managed object. GDMO is organized into templates, which are standard formats used in the definition of a particular aspect of the object, with rules for how these templates refer to each other. A complete object definition is a combination of interrelated templates. There are nine of these templates: Managed object class, Package, Behavior, Attribute, Attribute group, Action, Notification, Parameter and Name binding [46].

GDMO Standard Extension

The elements that at the moment form the GDMO standard do not make a reference to the knowledge base of an expert system. To answer these questions, it will be necessary to make changes on the template of the GDMO standard. In this section it is described how to accommodate the intelligent management requirements. To achieve this goal it was used the new extension GDMO+. This extension presents a new element RULE, which defines the knowledge base of the management expert system. This template groups the knowledge base supplied by an expert in a specific management dominion. It allows the storage of the management knowledge in the definition of the resources that form the system to be managed. The standard which was study contained the singular template RULE and its relations to other templates, Figure 5.

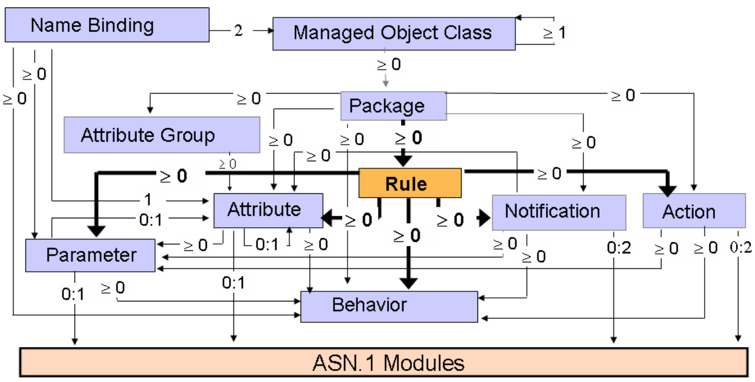

Figure 5: Relations between GDMO+ standard templates

Two relationships are essential for the inclusion of knowledge in the component definition of the network: Managed Object Class and Package template. In the standard GDMO+, both templates have the new property RULES.

This template is used to define the different kinds of objects that exist in the system. The definition of a managed Object Class is made uniformly in the standard template, eliminating the confusion that may result when different persons define objects of different forms. This way it is ensured that the classes and the management expert rules defined in system A can be easily interpreted in system B [47].

```
<class-label> MANAGED OBJECT CLASS
   [DERIVED FROM  <class-label>   [,<class-label>]*;]
   [CHARACTERIZED BY <package-label> [,<package-label]*;]
   [CONDITIONAL PACKAGES
           <package-label>  PRESENT IF condition;
          ,<package-label>] PRESENT IF condition]*;]
REGISTERED AS object-identifier;
```

DERIVED FROM plays a very important role when determining the inheritance relations which make it possible to reutilize specific characteristics in others classes of managed objects. In addition, a great advantage is the reusability of the object classes and therefore of the expert rules which are defined. This template also can contain packages and conditional packages, including the clauses CHARACTERIZED BY and CONDITIONAL PACKAGES.

This template is used to define a package that contains a combination of many characteristics of a managed object class: behaviors, attributes, groups of attributes, operations, notifications, parameters, behavior and notifications. In addition to these properties it was added a new property called RULES, which contains all the specifications of the knowledge base for the expert system [48]. Next definition shows the elements of a package template, in which it is possible to observe the new property RULES.

```
<package-label> PACKAGE
   [BEHAVIOUR <behaviour-label> [,<behaviour-label>]*;]
   [ATTRIBUTES <attribute-label> propertylist [,<parameter-label>]*
                  [,<attribute-label> propertylist [,<parameter-label>]*]*;]
   [ATTRIBUTE GROUPS   <group-label> [<attribute-label>]*
                                [<group-label> [<attribute-label>]*]* ;]
   [ACTIONS <action-label> [<parameter-label>]*
              [<action-label> [<parameter-label>]*]* ;
   [NOTIFICATIONS <notification-label> [<parameter-label>]*
                     [<notification-label> [<parameter-label>]*]* ;]
    [RULES     <rule-label>   [,<rule-label>]*;]
REGISTERED AS object-identifier;
```

All the properties which were defined in the package would be included later in the Managed Object Class Template, where the package is incorporated. Furthermore a same package can be referenced by more than one class of managed objects. Like the rest of the other properties defined in a package, the property RULES need a corresponding associated template.

<u>Expert Rule Template</u>

There are a number of different knowledge representation techniques for structuring knowledge in an expert system. The three most widely used techniques are expert rules, semantic nets and frames [49]. This system was design with expert rules. Knowledge is represented in production rules or simply rules. Rules are expressed as IF-THEN statements which are relatively simple, very powerful as well as very natural to represent expert knowledge. A major feature of a rule-based system is its modularity and modifiability which allow for incremental improvement and fine tuning of the system with virtually no degradation of performance [50].

In the standard GDMO+ the template RULE permits the normalised definition of the specifications of the expert rule to which it is related. This template allows a particular managed object class to have properties that provide a normalised knowledge of a management dominion. The structure of the RULE template is shown here:

```
<rule-label> RULE
   [PRIORITY       <priority> ;]
   [BEHAVIOUR   <behaviour-label> [,<behaviour-label>]*;]
   [IF          occurred-event-pattern [,occurred-event-pattern]*]
   [THEN      sentence [, sentence]* ;]
REGISTERED AS object-identifier;
```

The first element in a template definition is headed. It consists of two sections:

- <rule-label>: This is the name of the management expert rule. Rule definitions must have a unique characterizing name.

- RULE: A key word indicates the type of template, in our case a definition template and the specifications for the management expert rule.

After the head the following elements compose a normalized definition of an expert rule:

- BEHAVIOUR: This construct is used to extend the semantics of previously defined templates. It describes the behavior of the rule. This element is common to the others templates of the GDMO standard.

- PRIORITY: This represents the priority of the rule, that is, the order in which competing rules will be executed.

- IF: It contains all the events that must be true to activate a rule. Those events must be defined in the Notification template. The occurrence of these events is necessary for the activation of the rule and the execution of their associated actions. It is possible to add a logical condition that will be applied on the events occurred or their parameters.

- THEN: This gives details of the operations performed when the rule is executed. These are actions and diagnoses that the management platform makes as an answer to network events occurred. Those operations must be previously defined in the Action template.

- REGISTERED AS is an object-identifier: A clause identifies the location of the expert rule on the ISO Registration Tree. The identifier is compulsory.

Case of study

To show the viability of our proposal, it was carried out the study and building of a management expert system, so that the corresponding knowledge base begins to belong to the normalized proprieties information defined by the managed resources. It was provided a rule-based expert system applied to the fault diagnosis in telecommunication system of a power utility. The communications systems employed to implement the integrated intelligent management prototype belongs to the SEVILLANA-ENDESA a major Spanish power utility, Figure 6.

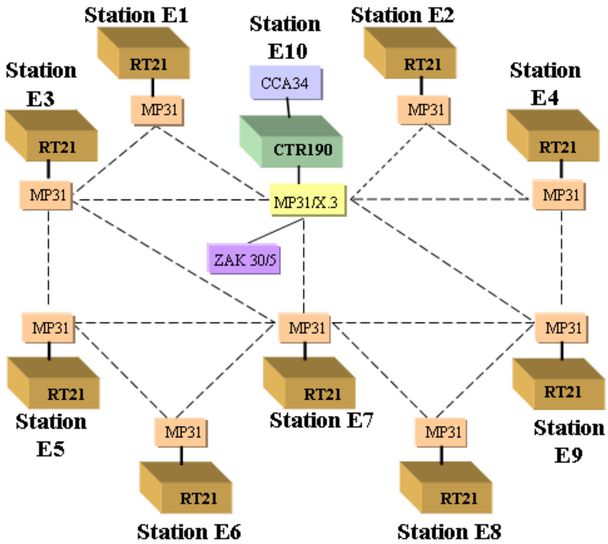

Figure 6: Power Company Network

Before management and control of that network was based on another expert system called NOMOS also developed by the Electronic Technology Department in the University of Seville [51]. This tool understands transceivers and multiplex equipment. Afterwards it was presented an extension of this expert system that integrated the knowledge base and the resources definition into a unique specification. The characteristic of the employed resources and the expert knowledge base use unique specifications. The architecture of this expert management system is based on the principle of the standard GDMO+. To define these specifications, the syntax and rules investigated in GDMO+ standard are used. The knowledge base of this system is integrated in the specifications of the resources using for that purpose our GDMO+ proposal. These new specifications contain management information of managed resources and include also the set of expert rules that provides the knowledge base of the expert system.

The following example is an expert rule named transmissionError, take charge of detecting failures in the data transmission module of the transceiver equipment.

```
transmissionError RULE
  PRIORITY 4;
  BEHAVIOUR transmissionErrorBehaviour;
  IF (?date ?time1 ?local 7_TX_C2 ?remote ALARM)
     (?date ?time2 ?local 7_TX_C2 ?remote ALARM & : (<(ABS(? ?time1 ?time2)) 1.00))
 THEN (“Severity:" PRIORITY) ,
```

("Diagnostic: "   It damages in the modulate transmission", ?local),
("Recommendation "Revision transceiver");
REGISTERED AS {nm-rule 2);

<u>AN EXPERT SYSTEM FOR THE DESIGN OF A HFC TELECOMMUNICATION NETWORKS: DATACAB</u>

<u>Introduction</u>

A typical cable network, which covers a medium-sized city, consists of an HFC (Hybrid Fiber Coax) network which is made up of an optical fiber network and its continuation into a coax network. Potentially, a parallel telephony network is deployed [52][53].

The cable network structure supports cable television, internet access and telephony services and it is represented in Figure 7. The structure of the network depends on each country and operator, but typically a fiber optic network is made up of a head-end connected to a primary network, which contains a number of primary nodes, for example servicing between 30.000 and 90.000 homes each. These main nodes are connected to a network of several secondary nodes (secondary networks), each of which typically serves around 10.000 homes. In general, the secondary nodes also contain an RTC (Remote Telephony Center). The tertiary networks (distribution networks) connect to the secondary node and end in ONTs (Optical Network Terminations), which are connected to a Coax Network. This Coax Network finally distributes the signal to the subscribers.

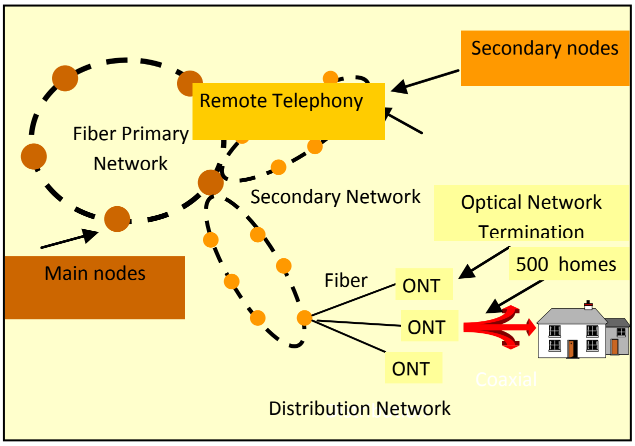

Figure 7: The HFC network to be designed

In the design of a HFC network the layout of the first, secondary and distribution fiber networks are normally easy to carry out because the location of their elements depends on the availability of adequate premises. Thus, the part of the network, which particularly requires a deep expert knowledge during design, is the coax network.

When a human designer carries out a coax network design he in many cases is only guided by his expert knowledge and intuition, and very basic support tools. The designer works this way: with the help of a specific software which reads and represents the geographical information from a GIS he begins the layout at the ONT (Optical Network Termination) progressively placing the cables and elements and carrying out the calculation of the signal level. If at some point it is not possible to reach the desired level signal on a particular connection with the layout in process, the designer has to redesign all or a part of the layout. Following this commonly used approach, the average time of producing a finished design for an ONT zone of around 500 connections is around 8 hours.

<u>Objective</u>

Datacab is the name of an expert system which was designed and implemented based on rules for the automatic design of a HFC telecommunication network and which works in an integrated manner with a GIS.

The great advantage of the proposed expert system approach is the great reduction of human workforce needed for the design of HFC network. With Datacab, instead of doing the designs himself, the human expert only has to verify the different designs generated by Datacab. Another important advantage is its adaptability, being possible to easily modify the knowledge base if the design criteria changes, and its high integration, working as an add-on software to the existing GIS. Finally, similar to other expert systems, the developed system can be used to train new staff.

The GIS into which integrated Datacab is GE SmallWorld [54], a powerful commercial geographical information system with special deployment in the areas of distribution networks like water, electricity, gas and telecommunication networks.

Datacab performs the design of the HFC network through the phases represented in Figure 8.

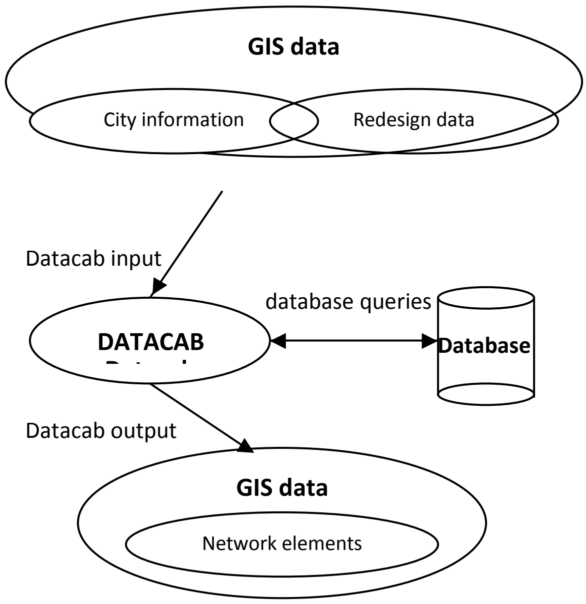

Figure 8: Datacab's interaction with GIS

- Phase 1: Datacab processes the geographical information corresponding to the zone to be designed, adding it to its knowledge base. The input data of the different elements of the zone with their location and all redesign information (i.e. possible access to accommodations, previous channel locations, non-permitted façades, etc.) is obtained from the GIS.

- Phase 2: By means of a graph representation of the GIS elements of the zone and the mechanisms of application of the knowledge, Datacab carries out the necessary calculations to achieve an optimal solution for the cable network design. Previously, a database access is realized in order to obtain the different parameters, such as the degree of attenuation of the cables or the gain of the available amplifiers for the different commercial network elements typically installed by the HFC operator.

- Phase 3: Datacab processes the obtained information, generates the network design based on the rule database and provides the GIS with the geographical information of the elements belonging to the cable network.

Rule-based system

When the geographic information is read from the GIS, each element of the zone, like a street or a building, is converted into an object of the class diagram and a connected graph is generated. The designed algorithm was based on Kruskal's algorithm [55], which selects the edges sorted by cost avoiding cycles in the layout. Kruskal's algorithm selects the lowest cost edge in each one of the iterations, as long as it does not generate any cycles.

Note that when a human expert carries out the design of a network, the radio frequency elements are placed at the same time as the layout is determined, which resembles more the second algorithm. However, this approach was discarded because of its higher system requirements and because the first algorithm proved to result in similarly good design results.

The difference between Kruskal's algorithm and our algorithm lies in the dynamic characteristics of the latter one. While the edge cost in Kruskal's algorithm does not depend on the previously selected edges, in our algorithm the cost will be allowed to vary.

It is distinguished two kinds of edge cost in the algorithm:

- Static cost: the static cost of a certain edge is defined as the cost which does not depend on the previously selected edges for the layout by the algorithm, i.e., the cost which is independent of the state of the layout such as the longitude of a certain street and the necessary cable to cover it.

- Dynamic cost: the dynamic cost is defined as the cost which at same point depends on the previously selected edges for the layout; i.e., the cost which is related to the state of the layout as, for example, the cost of a crossing between a placed cable and another cable (the company tries to avoid such cases).

On the other hand, the previous costs are determined by a priority number which was defined for the firing of each one of the rules. Thus, if the firing of two rules at same point was possible, the one with the higher priority would be fired and therefore it would get assigned a lowest cost.

Once the layout is achieved, the network elements are placed by means of a group of rules which calculates the signal levels corresponding to each one of the nodes of the layout. They thus determine the elements of amplification due to the calculated attenuation in order to provide the subscriber with the right signal levels.

These rules call methods belonging to the different radio frequency objects which are derived from the class radio-frequency. There are two especially useful methods of signal level calculation:

- Regressive signal-level calculation. This method calculates regressively the level of signal at a certain object. Therefore, it calculates the signal level belonging to an element and all the elements connected from it until the ONT.

- Progressive signal-level calculation. This method calculates progressively the level of signal at an object. Thus, it calculates the signal level belonging to an element and all the elements connected from it until the different points of connection in which it is involved.

The knowledge data is represented in the form of if-then rules, or, to put it another way, of rules made up of premises and conclusions to which a salience mechanism was added in order to provide them with a priority order. This system of reasoning has been chosen instead of, for example, other kinds of reasoning, such as one based on cases, because the design of an HFC network is based on rules and on the experience of the experts.

There are five main groups of rules in our system:

- Rules of initialization (R1): a group of rules which carry out the pre-processing of the information of the streets from the GIS, converting this data into DataCab information. In addition, they carry out a series of previously necessary calculations for the execution of the algorithm.

- Rules of sorting of edges by static cost (R2): a group of rules which carries out a sorting by static cost of the different streets of the zone. This group of rules comprises the rules for the division of zones, e.g., from 500 homes into 4 zones of about 125 homes.

- Rules of sorting by dynamic cost and selection of edges (R3): these rules sort by dynamic cost the different streets of the zone. In addition, this group contains the necessary rules to select from the previous order the edges necessary for the layout.

- Rules of location of elements (R4): a group of rules belonging to the placement of the different necessary radio frequency elements for the correct wiring of the zone. By firing these rules, which calculate the signal level at each node, the different RF elements (i.e. amplifiers, closets, etc) are located. If it is not possible to locate the elements, it is necessary to go back and fire the next group of rules.

- Rules of generation of the output data toward the GIS (R5): these rules generate the files with the necessary GIS information for the specification of the different elements belonging to the resulting layout, as well as their corresponding parameters.

Firstly, Datacab is initialized and the GIS data are read by means of firing the group R1 of rules. Later, the ordering of the several edges is carried out by means of the groups R2 and R3 of rules. The ordering from dynamic sorting generates a layout. The elements are placed and the signal calculations are carried out by means of the R3 group. If it is possible to reach the desired level signal with the generated layout from the previous phase, the design is considered good and the GIS data generated. Otherwise, another static order is forced and therefore another layout designed.

The total number of rules is 147. Table 1 shows some examples of rules for each group. The table contains 4 columns with information for each of the rules. The first column indicates the phase in which the rule is fired, the second one contains the name of the rule; the third one shows the priority in the firing, with a lower number meaning a lower priority, and, finally, in the last column a brief description of the rule is given.

Table 1: Some rules of each group

PH.	RULE	P.	DESCRIPTION
R1	CONV1	900	Converts the facts obtained in the reading of the file into Datacab objects
R1	CONV2	800	Completes the attributes of the objects created by rule CONV1.
R1	R1-R2	-1	Carries out the additional steps from phase R1 to phase R2.
R2	ONT	900	Selects the edges which reach the ONT.
R2	FAÇ	700	Selects the edges which represent cable preferably over façade versus cable through tubes.
R2	LNG	100	Sorts the edges by length of cable.
R2	DIV	900	Separates the streets belonging to each zone of 125 homes.
R3	CIC	900	Detects cycles in the layout.
R3	AVN	800	Selects edges belonging to main streets.
R3	PRN	900	Pruning of the non-necessary edges for the layout.
R3	ADJ	600	Selects the edges adjacent to the currently selected edges.
R4	CABZ1	900	Places the wiring of the first 125-home zone.
R4	CONN2	700	Places a splitter with one input port and three output ports.
R4	AMP	600	Places the necessary amplifiers for the layout.
R4	R4-R5	-1	Carries out the step from phase R4 into phase R5.
R5	CONV	900	Converts the objects generated by Datacab into GIS information

<u>An Example Use</u>

For this example, a zone of the city of Seville, Spain, was selected. More precisely, a node which is placed at Llerena Street, in the center zone of Seville and which connects 370 points.

Thus, the design of a zone by means of Datacab includes the follow stages:

1. A technician obtains the re-design information of the zone by means of a visit to the zone, which must include e.g. possible access to accommodations, previous channel locations, non-permitted façades, etc. The duration of this stage is 2 hours.

2. The previous information is introduced into the GIS. The duration of this stage is around 3 hours.

3. The GIS generates the gis.out file. The duration of the stage is 28 seconds.

4. Datacab reads the gis.out file and obtains the gis.in file (with the design of the zone) in 1 hour and 23 minutes by means of the execution of the following stages:

    - For each element read from the gis.out file, an object is generated.

    - The properties of the object related to the technical characteristics of the different elements for the design (such as the degree of attenuation of the cables or the gain of amplifiers) are filled in from the database of elements.

    - Datacab fires the 5 groups of rules (R1-5); 125 rules are fired and the result is a design which includes 350 design elements. The cable structure of the design is shown in Figure 9.

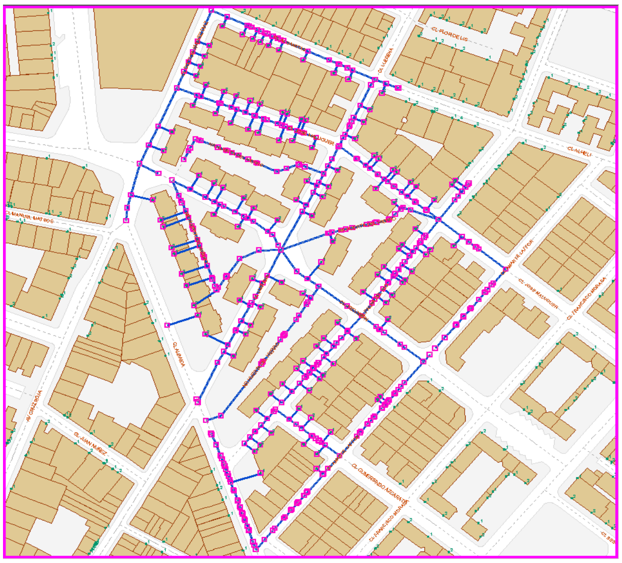

Figure 9: A layout generated by Datacab

5. The GIS reads the generated gis.in file relating to the design and stores the corresponding information. The duration of this process is 1 minute and 20 seconds.

6. The design is validated and approved by a human expert. The time of this validation is 42 minutes.

This zone was also designed by a human expert without the use of Datacab. In the human design, the stages 2, 3, 5 and 6 were not considered and the stage 4 was replaced by a human design which took around 7 hours.

Therefore, once calculated the total design time for both cases, the total time took by means of Datacab for all design phases was around 7 hours as opposed to around 10 hours of the human expert which means a 30% time saving. Besides in the Datacab design the above-mentioned stage 2 made it possible to store the re-designing information in GIS for future uses. The deployment costs of the elements like cable meters, splitters, amplifiers...etc placed by Datacab and the human design were similar.

AN EXPERT SYSTEM FOR THE DIAGNOSIS OF DIGESTIVE DISORDERS: SEDDIC

Introduction

Diarrhea affects all ages and races, and nowadays it is an important cause of morbidity and mortality worldwide. Infectious diarrhea is the main cause of infant mortality in developing countries. Each year more than 5 million children in the first year of life die for this cause. Diarrheal diseases are also a considerable socioeconomic problem causing great absenteeism.

Diarrhea is a sign that reveals a pathophysiological alteration of one or more functions of the intestine (secretion, digestion, absorption and motility) and which ultimately indicates a disorder of intestinal transport of water and electrolytes. Diarrhea is defined as increased volume, fluidity or frequency of bowel movements in relation to bowel habits of each individual. Since 60-75% of stool weight for the water, diarrhea is mainly due to excessive fecal water.

Chronic diarrhea is a situation in which there is a greater fecal excretion of 200 grams per day for a period exceeding 2 weeks.

With the aim of diagnosing diarrhea, doctors begin with a diagnostic approach aimed to rule out causes of easy diagnosis or, at least, to guide the diarrheal syndromic and identified three possible starting points: steatorrhea (excessive fat loss in the faeces), watery diarrhea, or a colonic disease.

Objective

There are hundreds of cases that can cause chronic diarrhea (from a poorly administered medication to a possible tumor). A large number of them is easily diagnosed and require few additional tests. But, on the other hand, it is extremely difficult to know the exact etiology in a substantial number of cases. This causes that medical professionals encounter many problems in order to make a diagnosis of high reliability.

Among all the documentation the reference point is situated on the main scheme (Figure 10) of the disease developed by Dr. Laso in his book: "Differential Diagnosis in Internal Medicine." [1].

SEDDIC is an expert system which aims to help the specialist in digestive diseases in the diagnosis of chronic diarrhea.

Rule-based system

The previously shown scheme covered perfectly the needs of our expert system, in which the facts would store the most of the symptoms and patient data, firing them a set of rules for the increased likelihood of some diagnoses (there is a maximum recall saying that no medical diagnosis can be considered 100% accurate) or the discarding of others.

Thus, the knowledge acquisition was structured in a sequential way, trying to imitate the behavior of the physician in his/her medical consultation. This procedure distinguishes the following phases:

- Collection of symptoms and physical examination: In this first phase, the expert system asks about the existence or not of different symptoms related to characteristics of diarrhea syndrome (frequent stools, diarrhea processing time, volume of faeces, blood or pus in the Lee ... etc.) and the outcome of the physical examination of the patient carried out by the doctor.

- Collection of early diagnosis: SEDDIC processes the results of the previous step (patient data, symptoms and physical examination), firing a first set of production rules which obtains a number of possible diagnoses. The result of this phase is a set of diagnoses along with a corresponding degree of safety. This degree of certainty or probability ranges from 0 to 5, depending on the number of symptoms associated with a disease suffered by the patient.

- Grouping of diagnoses by symptoms: The work of restructuring carried out internally by SEDDIC aims to bring together, from the previously obtained set of diagnoses, those that use the common symptoms. Once carried out this task the knowledge is prepared to be displayed to the user.

- Display of preliminary results: The system shows the first conclusions as possible diagnoses. In the case that anyone of the diagnoses is of a high reliability (4 or 5 of probability in the designed scale), the doctor is informed and given the option to terminate consultation or continue with the same process of diagnosis oriented to other alternative.

- First analysis: At this point, it is requested the results of a series of tests and additional examinations on the patient, including blood, urine, stool, and rectal examination. If the patient has not these results, the consultation will continue with it later.

- Obtaining the second diagnosis: Using data from the analysis requested by SEDDIC, it is generated a set of additional diagnoses, following the same procedure as in the first phase in the case that some diagnoses were of high reliability.

- Decision on the diagnostic tree: Once you get the result of the consultation on this point, without a clear diagnosis or without it being accepted by the physician, a decision is taken for any of the three main routes of diagnosis: steatorrhea, diarrhea of colonic origin or chronic watery diarrhea. Once chosen a diagnosis (the selection will be calculated with the percentage of symptoms suffered by patients in each one) and checking that it was not taken before, the system will trigger the analysis to that route.

- Route of steatorrheic, colonic or aqueous diarrhea: Once the diagnosis is oriented towards one of these three cases, the systems turns on to their analysis and, following the guidance and knowledge relative to the treatment of this disease, it request the results of other analysis and the corresponding route through the feature tree.

- Route of dark diarrheic process: In the case that SEDDIC arrives to this phase and it has not a main diagnosis, it will be discussed a possible dark diarrhea and the possible conclusions are obtained with the set of all data.

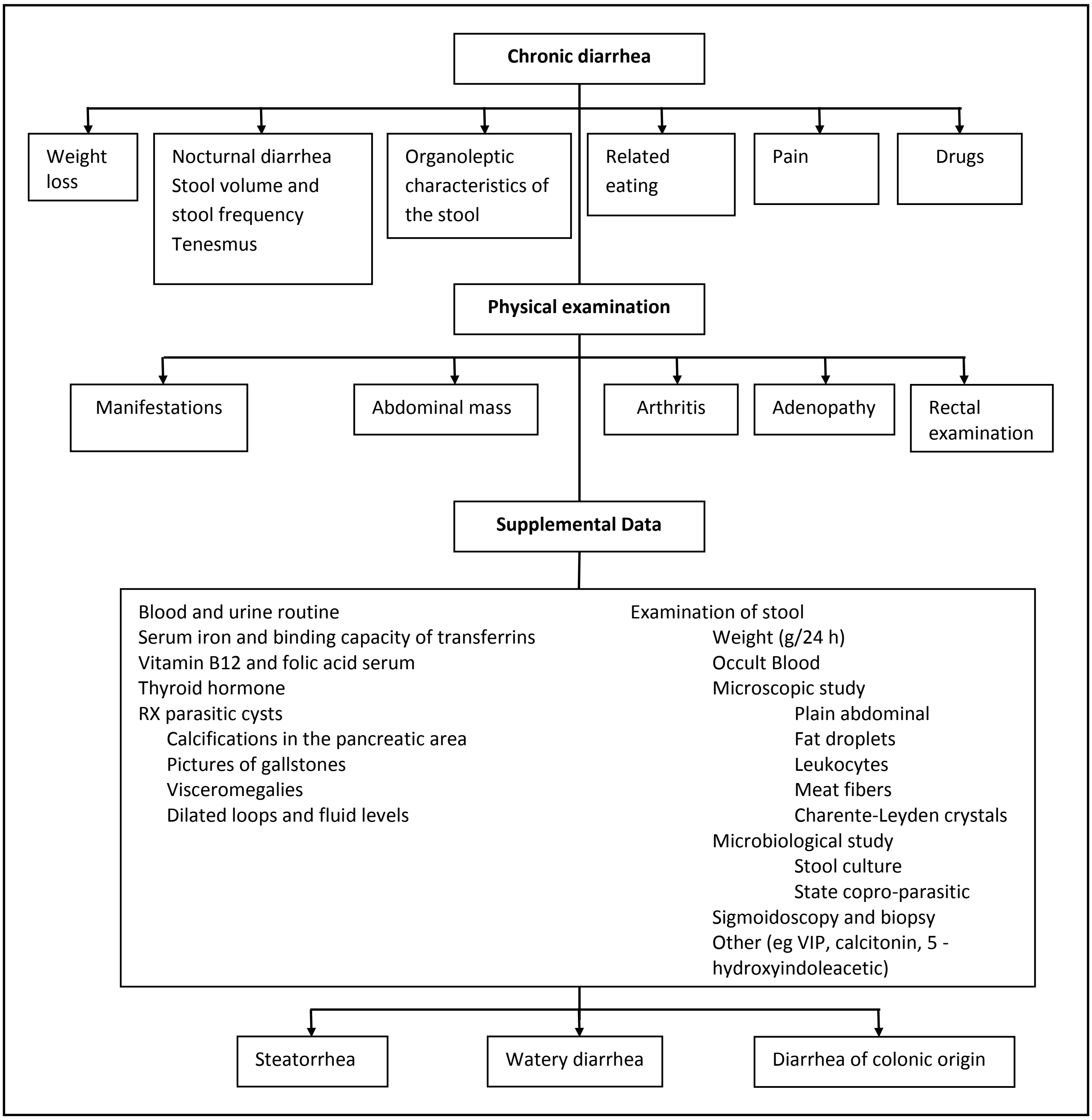

Figure 10: Initial evaluation of chronic diarrhea

Finally, the language CLIPS was later used for programming tasks and to generate the system. Later, the tests for verification and validation of the system were carried out.

Tests and conclusions

The first of two essential tasks in the assessment of any expert system was carried out by conducting multiple tests (see Table 2) in which abnormalities were corrected in the knowledge kernel of SEDDIC.

Table 2: Verification of the system

SYMPTONS	TEST 1	TEST 2	TEST 3
Duration	14	7	21
Frequent bowel movements	Yes	No	No
Weight loss	No	Yes	No
Stool volume	Normal	Abundant	Abundant
Appearance of stools	Paste	Solid	Liquid
Blood	No	No	No
Mucus or pus	Yes	Yes	No
Nocturnal diarrhea	No	Yes	No
Laxative Drugs	No	No	No
Diarrhea after eating	No	Yes	No
Pain	Yes	No	No
Tenesmus	No	Yes	No
Hyperpigmentation	Yes	No	No
Telangiectasias	No	Yes	Yes
Eczematous lesions look	No	Yes	No
Dermatitis herpetiformis	Yes	No	No
Oral thrush	No	Yes	Yes
Pioestomatitis vegetans	No	Yes	No
Lower extremity nodules	Yes	No	No
Septal panniculitis	No	Yes	Yes
Streptococcal infections	Yes	No	No
Pyoderma gangrenosum	No	Yes	No
Sweet Syndrome	No	Yes	Yes
Splenomegaly	No	Yes	No
Hepatomegaly	Yes	No	No
Left quadrant mass	Yes	No	No
Right quadrant mass	Yes	No	No
Arthritis	No	Yes	No
Lymphadenopathy	No	Yes	Yes
Fistulas	No	Yes	No
Perianal abscess	Yes	No	Yes

The validation of the system was carried out with the collaboration of Dr. Ovidio Belda Laguna of the Hospital Universitario Virgen del Rocío (Seville – Spain).

AN EXPERT SYSTEM FOR THE DETECTION OF FRAUDS IN A POWER SUPPLY

Introduction

Mainly, the utilities present two classes of incidents relative to the losses of energy:

- Technical losses. These losses are produced in the distribution stage. In the power distribution companies, they correspond with energy losses related with: wire warming (Joule Effect), distribution facility blemishes and natural reasons.

- Non-technical losses. Faults and/or manipulations on the installation that induce the total or partial absence or modification of the consumption on the company side. If the company cannot control the consumption correctly it is not possible to invoice the utility and, therefore, an economic loss is produced. The most common are:

  1. Anomaly. They are produced by breakdowns or mistakes by the company installation technical personnel or by deterioration of the equipment.
  2. Frauds. Manipulations illegally realized by clients in their installations, with the objective of modify the amount registered on the meter.

Currently, non technical losses cannot be forecasted easily. However, technical losses can be forecasted because utility distribution network is continuously monitored.

Utility distribution companies have a lot of information about clients' facilities and their utility consumption. In case of power utilities, there are a lot of customers in low voltage, because the domestic supplies normally have a low voltage. The control of these customers is realized by two methods:

- Telemeasuring. Usually, in customers with high power connection, the facility has a MODEM. It is used to transmit different information about customer's facility: consumption, phasors, etc. This information can be used to take a better power supply decision or to make a better detection of non-technical loss.
- Manual measuring. Normally, in customers with low contracted power, the information about facility is provided by a company employee, who gets around the client location and manually takes the information. This information can be collected monthly or bi-monthly. Company employees gather information about consumption, but, sometimes, they provide additional information about clients' facility using comments which can be added in the company databases.

In the same way as manual measuring, when inspectors and technical staff of the companies make customers' facility inspections, they can provide information about customers, facilities or problems using comments in natural language. These comments provide additional information about the customers' facilities and habits.

When inspectors find a fraud or an anomaly in the customers' facilities, they will store all information about it in the company databases. They could use pre-formatted information and comments in natural language.

The inspectors have to analyze a great volume of information to determine the problems in customers' facilities. This process can take between 5 and 15 minutes by customer, depending on the quantity of available information about customer.

Solution

The solution proposed to the described problem is a Hybrid Expert System to analyze all the information about customers. Mainly, the Hybrid Expert System architecture is based on:

- The nature of the information to process will determine what type of techniques or technologies are most adapted to analyze it, since the same methods for numeric and alphanumeric information cannot be used directly.
- The way the information is presented to the user. The activities that the employee develops, determines which is the most suitable information format.
- The company infrastructure. The system must be measured correctly to allow the implantation in the company infrastructure.

These conditions are the same for all utility companies, because they use very similar information and volume of clients.

The influence of these three factors is determined by a pre-processing phase of knowledge extraction, selecting the information that can contribute something in the classification and detection process of non-technical losses.

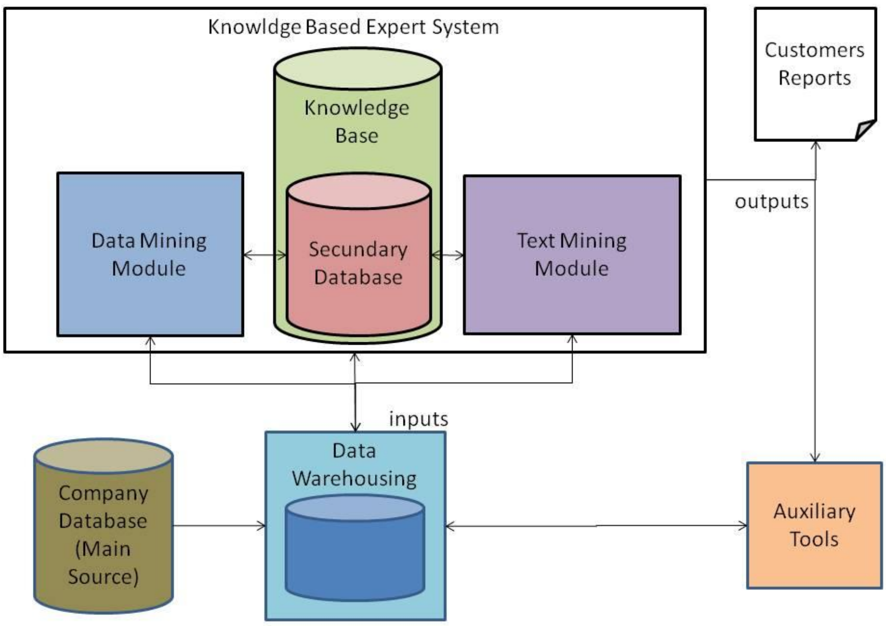

Figure 11: Architecture of the proposed Hybrid Expert System

The proposed architecture for Hybrid Expert System, showed in Figure 11, has the following modules:

- Data Warehousing Module. The company databases have a lot of information about customers': contracts, equipments, locations, customers' information, consumption data, inspectors' comments and documentation, etc. The system cannot directly work with all this information, because the system is busy with other processes (gathering information, registering new clients, etc.). In this way, the information has to be previously extracted from company databases. When the analysis is made on a lot of customers, the volume of information increases considerably. In this case, data warehousing is used to improve the state and quality of information, verifying the coherence, integrity and several kind of errors (format errors, incorrect values, etc.). In addition, this module provides information to other modules with the correct format. This module provides the inputs for the other modules.

- Data Mining Module. This technique generates a table of values which are used to make the utility consumption studies. The trends and ranges of consumption are established by means of statistical techniques. The studies of consumption ranges are carried out through the application of a statistic study which searches normal pattern behavior, but taking into consideration a series of criteria which allow the distinction of one type of consumption from the other. There are a number of key information fields for determining acceptable consumption standard patterns: geographical location, economic activity, billing frequency, time discrimination and contracted power (in other utilities as gas or water, volume flow). In addition, due to the different patterns that may be found, may be necessary to establish a temporary division in the information for consumption patterns: absolute, annual, seasonal and monthly. These groups provide with divisions within which are defined normal conditioned patterns by the client characteristics. This technique is used on a great sample to get groups with statistical significance. But it is necessary to pre-process the information: discretization of some values and filter customers with anomalous consumption. Additionally, the table which is obtained by this module allows making reports about general trends and statistics.
- Text Mining Module. This technique generates a dictionary which is used to analyze the documentation of inspections made in electrical installation clients. Initially, this module is made based on experience, using concept extraction processes on documentation and inspection comments of customers' facilities. These concepts are organized in several categories that identify several events in customer's facilities. In paper [1] a new method to increase the efficiency of concept classification process is described.
- Auxiliary Tools Module. This module makes the follow-up of the customers who presents certain characteristics in the analysis process.
- Rule-based Expert System Module. This is the main module. This module uses the rules generated by other modules and the expert rules to analyze all information about customers. Additionally, this module coordinates the results of the different modules and takes the decision about the customers' classification. The expert system uses some rule sets classified in seven different groups according to their function. In the power utility case, the system has 135 rules, but client analysis may apply more rules, adding the rules of other modules. There are some rules which have a dynamic antecedent which is filled with information of the data mining table module and the dictionary of the text mining module. There are a lot of rules because it is necessary to analyze all the information about the customers. In addition, this module gathers all

information necessary for inspectors and generates reports showing only the information required for inspections.

Conclusions

The proposed Hybrid Expert System researches focuses in area of automation of the available information analysis for non-technical losses classifications on utility companies. This system makes following contributions:

- Identification and classification of the necessary knowledge.
- This system reduces the time cost of the analysis. In the performed tests, the system approximately analyzed 14.000 customers in 16 hours.
- This system can be used as a support system for other frameworks. In this way, this system provides a complex filter that removes usually between 20% and 70% of customers who otherwise would be a wasted expenditure, since in most cases they do not have non-technical losses.
- The reports of the Hybrid Expert System provides to inspectors and researchers more information about the client and on the problems presented, as are just and necessary information, using graphic information on a temporary bases. Moreover, in these reports, the reasons why the client has been included in a given category can be found.
- Hybrid Expert System used as a Decision Support System (DSS) for utility companies, using reports and graphics and making the process of selection for inspection easier.
- Testing of the Hybrid Expert System in real cases related with power utility.

REFERENCES

[1] Liebowitz, J ed. "The handbook of Applied Expert Systems". CRC Press, 1998.

[2] ITU-T - Recommendation X..711 : Common Management Information Service for CCITT applications – 1991.

[3] Rabinowitz, H. Flamholz, J. Wolin, E., and Eucher, "J. NYNEX MAX: a telephone trouble screening expert". Proceedings of IAAI91, (pp. 217-230). Anaheim, California. 1991.

[4] Chen, C; Hollidge, T.; and Sharma, D.D. "Localization of troubles in telephone cable networks". Proceedings of AAAI96/IAA896, (pp. 1461-1470). Portland, Oregon. 1996.

[5] Raman, Lakshmi G., "Fundamentals of Telecommunications Network Management," NJ, IEEE Press, 1999.

[6] Liebowitz, J. "Expert System Applications to Telecommunications", John Wiley & Sons, Inc. New York. 1988.

[7] Ezawa, K; Norton, S., "Constructing Bayesian networks to predict uncollectible telecommunications accounts". Journal of IEEE Expert, Vol. 11, No. 5. (pp. 45-51). 1996.

[8] K. L. Lo, Sabir; S. Ghauri, "Expert System For The Design Of Distribution Networks", IEEE International Conference on Advances in Power System Control, Operation and Management, November 1991, Hong Kong.

[9] Nawal A; El-fishawy, Salah A Khamis, "An Expert System Architecture for the Design of Local Area Networks", Vehicular Technology Conference, 2000. IEEE VTS-Fall VTC 2000. 52nd , Volume: 4, 2000.

[10] Chih-Hung Wu; Shie-Jue Lee, "An Object-oriented Expert System for Local Area Network Design", Computing and Information, 1993. Proceedings ICCI '93., Fifth International Conference, 1993.
[11] Shurkin J., "Medical Applications. Al moves from the lab to the doctor's office". IEEE Expert 1986: 1. 10-12.
[12] Kulikowsky C., "Artificial intelligence in medical consultation systems: A review", IEEE Engineering in Medicine and Biology Magazine 1988; 7. 34-88.
[13] Szolovits P, Patil As, Schwartz Wb., "Artificial intelligence in medical diagnosis". Ann Int Med 1988; 108: 80-87.
[14] Kar A, Millea Ge, Shepard Sv. "Pulmonologist: a computer-based diagnosis system for pulmonary diseases", Int J Bio-Medical Computing 1987; 21: 223-235.
[15] C.A.Perry, "Knowledge Bases in Medicine: a Review", Bull Med Libr Assoc 78(3) July1992. Pp. 271-282.
[16] Shortliffe, E.H, "Medical Expert Systems: Knowledge Tools for Physicians", Western Journal of Medicine, 145, 1986, 830-839.
[17] Liu, Chen-Ching, Dillon, Tharam, "State-Of-The-Art Of Expert System Applications To Power Systems", Electrical Power & Energy Systems Vol. 14 No. 2/3 April/June 1992, Pp. 86-96.
[18] Germond, A.J., Neibur, D. "Survey Of Knowledge-Based Systems In Power Systems: Europe", Proceedings Of The Ieee, Vol. 80, No. 5, May 1992, Pp. 732-744.
[19] Behera, H.S.; Dash, P.K.; Biswal, B., "Power Quality Time Series Data Mining Using S-Transform And Fuzzy Expert System", Applied Soft Computing 10 (2010) Pp. 945-955.
[20] Garcia, Antonio; Biscarri, Félix; León, Carlos; Personal, Enrique, "Estimation Of The Line-To-Line Voltage Phases In Mv Distribution Faults Using Transformer Inrush Currents", 1010 Ieee Power And Energy Society General Meeting. 25-29 July 2010, Minneapolis, Mn, Usa, Pp. 1-5.
[21] Sharifian, Hoda; Abayaneh, H. Askarian; Salman, Salman K.; Mohammadi, Reza; Razavi, Farzad, "Determination Of The Minimum Break Point Set Using Expert System And Genetic Algorithm", Ieee Transactions On Power Delivery, Vol. 25, No. 3, Pp. 1284-1295, July 2010.
[22] Abdelaziz, A.Y.; Mekhamer, S.F.; Nada, M.H. "A Fuzzy Expert System For Loss Reduction And Voltage Control In Radial Distribution Systems", Electric Power System Research 80 (2010), Pp. 893-897.
[23] El-Werfelli, M.; Dunn, Rod; Iravani, Pejman. "Backbone-Network Reconfiguration For Power System Restoration Using Genetic Algorithm And Expert System", International Conference On Sustainable Power Generation And Supply, 2009. Supergen'09, Pp. 1-6, 6-7 April 2009.
[24] Burt, G.M.; Mcdonald, J.R.; King, A.G.; Spiller, J.; Brooke, D.; Samwell, R. "Intelligent On-Line Decision Support For Distribution System Control And Operation", Ieee Transactions On Power Systems, Vol. 10, No. 4, Pp. 1820-1827. Nov. 1995.
[25] Yujin, Li; Qing, Zhang; Deyi, Zheng; Pengyun, Xu; Baozhong, Wu. "On Operation Ticket Expert System For A Substation Using Object Oriented Technique", Proceedings Of The 27th Chinese Control Conference, Pp.2-5. Kunning, Yunnan, China. 16-18 July 2008.
[26] Piga, Erika; Geschiere, Alex. "Self Learning Expert System (Sles) For Power Transformers", 20th International Conference On Electricity Distribution, Cired, Prague, 8-11 June 2009.

[27] Yu, Xiaodong; Wang, Kui, ”Digital System For Detection And Classification Of Power Quality Disturbance”, Power And Energy Engineering Conference, 2009. Appeec 2009, Pp. 1-4. Asia-Pacific, 27-31 March 2009.
[28] León, C.; Mejías, M.; Luque, J.; Gonzalo, F. “Expert System For The Integrated Management Of A Power Utility’s Communication System”, Ieee Transactions On Power Delivery, Vol. 14, No. 4, Pp. 1208-1212. October 1999.
[29] Luque, Joaquín; Gonzalo, Fernando. “Integrating Power Utility Telecommunication Networks”, Ieee Computer Applications In Power, Pp. 27-30. April 1996.
[30] Qiu, Wangbiao; Qiu, Zhiyuan. “Design For Symmetrical Management Of Storage Battery Expert System Based On Single Battery”, Proceedings Of The 2006 Ieee International Conference On Mechatronics And Automation, Pp. 1141-1146, 25-28 June 2006.
[31] Cabral, J.; P. Pinto, J. O.; Gontijo, E.; Reis Filho, J. “Fraud Detection In Electrical Energy Consumers Using Rough Sets”, Ieee International Conference On Systems, Man And Cybernetics. (2004)
[32] Galván, J. R.; Elices, A.; Muñoz, A.; Czernichow, T.; Sanz-Bobi, M. “A. System For Detection Of Abnormalities And Fraud In Customer Consumption”, 12th Conference On The Electric Power Supply Industry. Pattaya, Thailand. November 2-6, (1998)
[33] Azadeh, A.; Ghaderi, S.F.; Tarverdian, S.; Saberi, M. “Integration Of Artificial Neural Networks And Genetic Algorithm To Predict Electrical Energy Consumption”, Applied Mathematics And Computation 186 (2007) 1731-1741.
[34] Azadeh, A.; Ghaderi, S.F.; Sohrabkhani, S. “Forecasting Electrical Consumption By Integration Of Neural Network, Time Series And Anova”, Applied Mathematics And Computation 186 (2007) 1753-1761.
[35] Nagi, J.; Mohammad, A.M.; Yap, K.S.; Tiong, J.K.; Ahmed, S.K. “Non-Technical Loss Analysis For Detection Of Electricity Theft Using Support Vector Machines”, 2nd Ieee International Conference On Power And Energy (Pecon 08), December 1-3, 2008, Johor Bauru, Malaysia.
[36] Nagi, J.; Yap, S,K.; Tiong, S.K.; Ahmed, S.K.; Mohammad, A.M. “Detection Of Abnormalities And Electricity Theft Using Genetic Support Vector Machines”, Tencon 2008, Ieee Region 10 Conference. 19-21 Nov. 2008, Pp. 1-6.
[37] P. Rohan, "Introduction to electromagnetic wave propagation", Artech House, London, UK, 1991
[38] National Communications System, Office of Technology and Standards,701 S. Court House Road Arlington, Virginia 22204-2198 , USA, 1996.
[39] R. Adair; D. Peach; D. Bodson, "The growing family of federal standards for HF Radio Automatic Link Establishment (ALE)", QEX July pp.3-8, ARRL, Newington, CT, USA, 1993.
[40] Eric E. Johnson; Robert I. Desordis Jr. et alt. "Advanced high frequency radio communication", Artech House, 685 Canton Street, Norwood, MA 020622, USA, 1997
[41] Radio waves propagation, ITU-R, P Series Recommendations, Radio communications Assembly, Geneva, Switzerland, 2010.
[42] W. Stallings, “SNMP, SNMPv2, and CMIP: the practical guide to network”, Publication Reading, Mass. Addison-Wesley, 2000.
[43] L. Goleniewski; K.W. Jarrett, “Telecommunications Essentials, Second Edition: The Complete Global Source”, Addison Wesley Professional, 2006.
[44] ISO/IEC DIS 10165-4 / ITU-T Recommendation X.722, Information Technology - Open Systems Interconnection - Structure of Management Information - Part 4: Guidelines for

the Definition of Managed Objects (GDMO), International Organization for Standardization and International Electrotechnical Committee, 1993.

[45] ITU-T, Recommendation M.3400, TMN Management Functions. Study Group IV, 1996.

[46] ISO/IEC and ITU-T, Information Processing Systems - Open Systems Interconnection - Systems Management Overview. Standard 10040-2, Recommendation X.701, 1998.

[47] N. Hebrawi, “GDMO, Object modelling and definition for network management”, Technology appraisals, 1995.

[48] J. Zuidweg, “Next generation intelligent networks”, Boston: Artech House, 2002.

[49] J. Giarratano; G.D. Riley, “Expert Systems: Principles and Programming”, Book, Brooks/Cole Publishing Co., 2005.

[50] R.J. Brachman; H. J. Levesque, “Representation and reasoning”, San Francisco, CA: Elsevier/Morgan Kaufmann, 2004.

[51] C. León; M. Mejias; J. Luque; F. Gonzalo, “Expert System for the Integrated Management of a Power Utility's Communication System”, IEEE Trans on Power Delivery, Vol. 14, No. 4, 1208-1212, 1999.

[52] S.Ovadia, “Broadband Cable TV Access Networks: From Technologies to Applications”, Prentice Hall, 1st. Edition, 2001.

[53] E.Tunmann, “Hybrid Fiber Optic/Coaxial (HFC) Networks”, CMP Books, 1st.Edition, 1995.

[54] P. Rigaux, M.O. Scholl, A. Voisard, “Spatial Databases: With Application to GIS”, Academic Press, 2002.

[55] R. Sedgewick, “Algorithms in C++”. Addison Wesley Iberoamerica, 1992.

[56] F.J. Laso, “Diagnóstico Diferencial en Medicina Interna” Harcourt Brace de España, S.A. 1997.