Application of machine learning techniques in water distribution networks assisted by domain experts

Luis M. Camarinha-Matos, Fernando J. Martinelli
New University of Lisbon and Uninova,
Faculty of Sciences and Technology
Quinta da Torre, 2825 Monte Caparica, Portugal.
cam@uninova.pt, fjm@uninova.pt

Abstract

This paper describes an ongoing work on the application of machine learning techniques in the domain of water distribution networks. This research is being done in the context of the European Esprit project Waternet. One part of this project is a learning system which intends to capture knowledge from historic information collected during the operation of a water distribution network. Captured knowledge is expected to contribute to improve the operation of the network. The ideas presented in this paper describe the first development phase of this learning system, focusing specially in the practical methodology adopted. The interaction between different classes of human experts and the learning system are discussed. Finally some preliminary experimental results are presented.

Keywords

Machine learning, water distribution network, knowledge acquisition, forecasting.

1 INTRODUCTION

Water. The importance of drinkable water in the human life is well known. The human beings are strongly dependent of the water and they tend to use much more water than the amount necessary to survive (Loucks & Costa, 1991). However in most cases people do not realize how difficult it is to send the water from its source to the taps at home with a good quality level. Instead, people become very demanding consumers, requiring from the water supply services very high levels of quality. Quality not only related to biological or physical-chemical factors, but also related to the continuity of the supply with adequate levels of pressure and flow.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35390-6_58

L. M. Camarinha-Matos et al. (eds.), *Intelligent Systems for Manufacturing*

Physical constraints. The complexity of the water distribution management is due to various factors including:

- The geographic and topologic structure of the networks. The total length of the pipes can typically reach 900 km, even in a normal water distribution system.
- Water distribution systems may be in permanent expansion or alteration due to population and economical modifications in the society.
- Sources of uncertainty like leakage due to accidental disruptions on the pipes or even from some pirate deviations.
- It is very difficult to find an adequate model to describe the behavior of the networks. Thus, the supervision and control strategies are based on manual pro manual procedures or heuristic rules.
- The heterogeneity of those systems. Each distribution network shows characteristics that are distinct from other ones.

Control constraints. The water distribution networks are operated from various pumping stations and reservoirs including some treatment stations (remote units). The control of these systems is in general performed locally and based on the operators' experience. Each station might have a control algorithm that handles routine situations. However, if any abnormal situation occurs, the operators are alerted and called to solve it, instead of the control algorithm. A central control station is normally available to supervise/coordinate the global network (central unit).

The control at the remote units' level is performed without any, or very little, coordination among the stations, loosing the overall view of the network. The absence of an overall view of the network makes difficult, costly and time consuming, the faults' identification and also their recovery. It is also difficult to find optimal strategies to operate the network.

Helping to find a solution. In the context of the European Esprit project WATERNET, an evolutionary knowledge capture approach for advanced supervision of water distribution is being developed. This project intends to develop a system to control and manage water distribution networks taking into consideration the necessary control of costs and required quality, and using the Information Technology to increase the level of automation and integration in those networks.

Field for machine learning. In some water distribution networks, there are large quantities of data collected during the operation of the network, which suggests the opportunity to apply learning techniques in order to find more optimized operation strategies. These data sets contain measurements of physical variables, water quality indicators, device status, operator actions and alarms reports.

Inside this project, this kind of data, captured from SMAS-Sintra a Portuguese water company, is being used to evaluate the application of machine learning techniques in the domain of water distribution network.

This paper focuses specially on the use of human guidance in all phases of the development and as a fundamental help element in the application of machine learning techniques to this domain.

Paper organization. The remaining of this paper is organized as follows. Section 2 introduces a brief summary of the WATERNET project, presenting the architecture with the subsystems composing the project and their interconnections. Section 3 introduces the Learning System, focusing the reasons for its development. Section 4 presents the work done to identify the areas where the machine learning techniques should be applied. Section 5 describes the difficulties detected in this work and the approach followed in the development of the system. Finally, section 6 presents some conclusions and open questions.

2 THE WATERNET PROJECT

The WATERNET project is a two year Esprit project that aims to design and develop an evolutionary knowledge and management system towards the control, decision support and optimal operation of drinkable water distribution networks. Included in the objectives of the project are the minimization of the costs of exploration, the guarantee of continuous supply of water with better quality monitoring, reduction of energy consumption and minimization of natural resources waste.

The system that is being developed in this project is composed by the following subsystems, as illustrated by Figure 1:

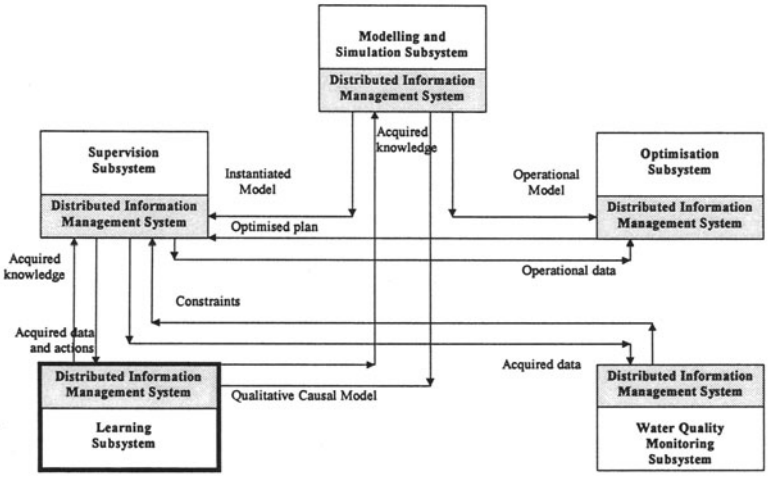


Figure 1 - WATERNET architecture

- A Supervision System, supervising the network in a distributed way, monitoring its current status, identifying deviations from the desired states and making decisions about the next control and management actions;
- A Distributed Information Management Subsystem (DIMS), supporting the cooperation and information exchange among sites and their activities (Afsarmanesh, Camarinha-Matos & Martinelli,1997);

- An Optimization Control Subsystem, producing optimized operational strategies for control devices (pumps, valves, ...) in order to minimize exploration costs (Quevedo et al. 1987);
- A Water Quality Monitoring Subsystem, monitoring the water quality in the network and guiding the water treatment process to guarantee the sanitary safety of the water;
- A Modeling and Simulation Subsystem, that is responsible for the network models, deriving the desired information for the other subsystems and making simulations about the network behavior;
- And finally, a Learning Subsystem, that is a system containing multiple learning algorithms, representing different paradigms, to support programming by demonstration and data mining on historical operation databases.

3 LEARNING SUBSYSTEM

The following aspects justify the inclusion of a learning subsystem in Waternet:

- The characteristic of continuous expansion and modification of the water distribution networks.
- The procedures used to operate the network, and the factors influencing the operation.
- The availability of historic data.

Continuous expansion. The water distribution networks are in continuous expansion or modification. Factors such as new housing or industrial zones request the network to grow in order to comply with the new supply needs. This growing process can be reflected in new pipes going to new regions or in increases in water demand, and will imply different strategies to operate the network. The introduction of a learning component represents a promising approach to cope with an environment with so many modifications along the time.

Operation. The operation of the network is performed locally in the remote units by control algorithms that deal with routine situations. However, when some abnormal situation happens, current algorithms alert the operator who becomes responsible to solve the problem.

In presence of an abnormal situation, i.e. when an alarm fires, the operator evaluates the status of the network and, based on his experience, decides which actions to perform in order to overcome the alarm. The reaction to an alarm can be as simple as a valve opening, but it can also be so drastic as the stop of one or various stations, or even the call of a brigade to fix any detected anomaly in the network.

This scenario suggests an interesting application case to machine learning techniques. If it is possible to find in the historic data various occurrences of one same status of the network during a specific alarm firing, and also the same actions performed by the operators to recover from this alarm, then it is possible to learn how to deal with it.

Historic data. The last factor, and probably the most important one, that supports the application of machine learning techniques in this project, is the availability of a large amount of historic data. The experiments described on this paper were performed over real historic data gathered in the SMAS-Sintra network. This historic data contains:

- Sensorial data (physical variables): All measurements of the physical parameters of the network are stored. This information represents flow and pressure values inside the pipes, reservoir levels and other measurements.
- Sensorial data (chemical variables): Measurements specifying the water quality are also stored, such as pH values and chlorine quantities.
- Device status: The status of some devices is also found in the stored data. Information if a pump is working or not and information about the percentage of opening of some valves are examples of these data.
- Operations: The operations performed by control algorithms or by the human operators, like the opening of valve or the turning on of a pump, are stored.
- Alarms: Every alarm reported in the stations is also stored.

These data is collected on a regular basis. In general, the readings of such information occur each 5 minutes. Data is also collected when any abnormal situation occurs. Each station of the network stores locally the data as text files. For instance, a file containing one day of collected data from a simple station can have sometimes 12 000 lines. SMAS-Sintra has 6 years of collected data for some stations. The complete amount of information available represents, in average, 4 years of 45 station amounts to something around 500 Mbytes of compressed data.

4 WHERE TO APPLY MACHINE LEARNING

4.1 Starting phase

The initial phase of this process was the identification and characterization of possible application areas for machine learning in the water distribution network domain. A strong interaction with the domain experts complemented with an analysis of the characteristics of the historic data led to the elaboration of a list of possible learning tasks. It is also important to notice that the objective is not to develop new learning algorithms but to apply existing results from the machine learning community. One of the main objectives is, therefore, to develop a practical working methodology that can lead to useful results.

The upper section of Figure 2 illustrates the use of these factors in the definition of a set of possible application areas for machine learning, or as stated there, a set of potential learning tasks in this domain.

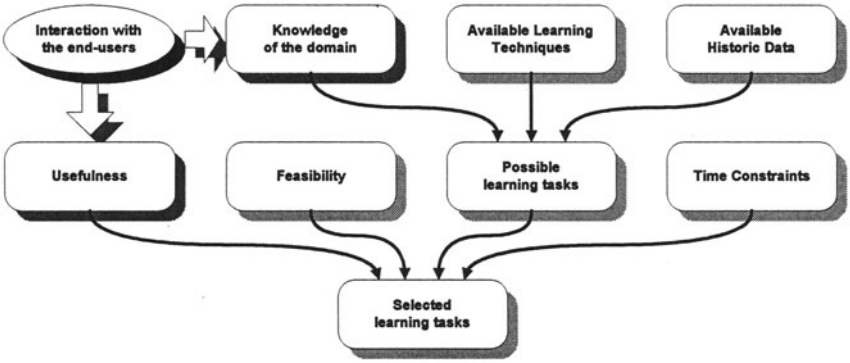


Figure 2 - Elements considered in the selection of the application areas for machine learning

4.2 Potential learning tasks

As a result of this preliminary work, a set of potential learning tasks in the water distribution domain was identified in (Afsarmanesh, Camarinha-Matos & Martinelli, 1997) and briefly summarized here:

a) Support for Production Planning: These tasks intend to extract some knowledge that can be useful in forecasting future water demand, and giving information on how the network is evolving.

b) Identification of factors that influence actions: These tasks have the objective to discover factors that influence the operation of some devices of the network in order to learn how to operate them.

c) Monitoring and alarm handling: These tasks aim to learn how to detect, diagnose and recover from failures observed in the network. The human operators' reactions to alarms will be used as training examples.

d) Preventive maintenance: These tasks will help the maintenance sectors of the water distribution companies in achieving a better management of their devices.

e) Improvement of user satisfaction: The evaluation of the user complains can bring some additional knowledge to the companies, especially in terms of anomalies detected in the network.

f) Other possibilities: Set of learning tasks that are related to the identification of the normal behavior of the network.

Around 16 learning tasks were identified. Of course this is not the complete set of possible application areas for this domain, but the identified ones illustrate the large potential for machine learning in water distribution systems.

4.3 Defining priorities

Due to the limited resources available and the short project duration, it was necessary to focus the work in a subset of the learning tasks presented above. In order to make a final selection, some additional factors were considered:

- Time Constraints: As the project was designed for 2 years, the idea to work on learning tasks that would require long time to finish was out of question. The work had to be focused on applications that would have some results by the end of the project.
- Feasibility: The feasibility of the learning tasks was also evaluated. Those tasks for which results could be more easily reached should be more strongly taken as a possibility.
- Usefulness: The factor with more weight in the decision was the usefulness of the application area. Priorities were defined according to the end-users' needs.

The domain experts were again of fundamental importance in this decision making process. Their judgement showed which application could lead, if successful, to higher added value to their work and which ones would be less important.

In order to facilitate the discussion with the domain experts, it was very important to perform some preliminary learning experiments showing their potential. The domain experts do not have basic knowledge of machine learning and some concrete examples are fundamental to help them understand what could be expected from these techniques.

This identification and selection phase was necessary because there is no documented tradition of practical application of machine learning in this domain. As a result of this process two learning tasks were selected for implementation:

- Water demand forecasting: This task is included in the group a) above. Its objective is to predict the water consumption in a region of the network in the near future. A good water demand forecast is very important to guarantee a continuous supply of water at a low costs
- Error monitoring and alarm handling: This task belongs to group c). The intention is to learn how the operators act in order to diagnose errors and recover from alarms occurring in the network.

5 HOW TO APPLY MACHINE LEARNING

5.1 Finding difficulties

Before developing the final software modules that support these learning tasks, it is important to perform a set of experiments in order to identify the main difficulties and decide on which algorithms to use.

As a result the following major problems were identified:

- Variables' selection: Some stations are characterized by hundreds of variables. Using all (or a large portion) of them in a learning activity is possible. The experiences have shown that a wide set of variables slows the learning process, due to the size of the search space. The results can also be inaccurate because some spurious information can led to wrong results.

- Pre-processing: The raw data available had to be pre-processed before they can be feed to the learning algorithms. This pre-processing has the objective to format the data in a way to focus the learning process. Sometimes high level features have to be extracted in order to represent, in a better way, the training objectives. For instance, the calculation of average values or derivatives, the transformation from flow values to consumed volumes, or even the mapping of variables' values for a different domain.
- Results assessment: Finally, the results obtained with the learning algorithms have to be evaluated in order to identify if they really have some physical significance that make possible their use in the supervision process.

Figure 3 illustrates these difficulties.

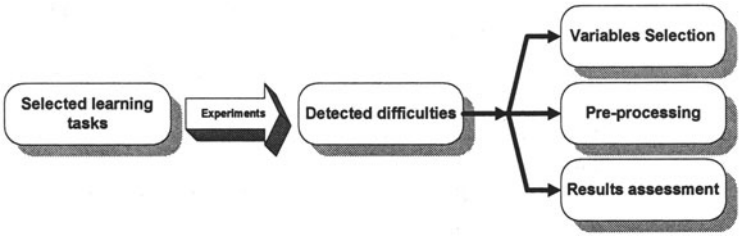


Figure 3 - Main difficulties detected in the learning process

The domain expert's knowledge represents an important help to overcome these difficulties, as illustrated in Figure 4.

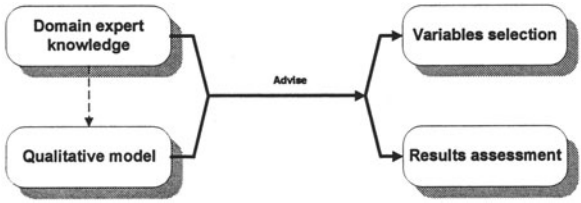
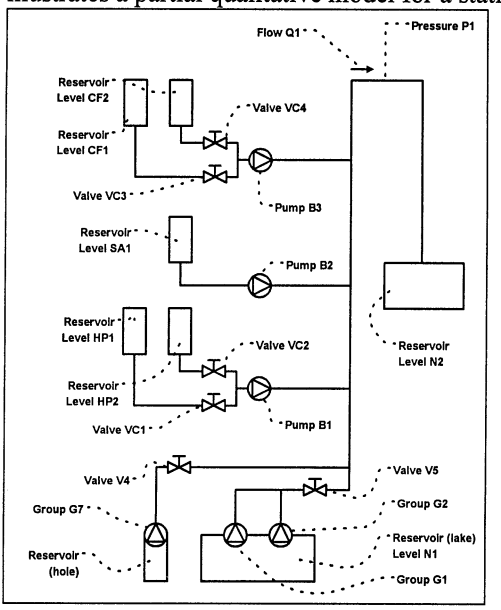


Figure 4 - Use of domain expert knowledge to help the learning process

The domain experts know the network and have some heuristic knowledge about the relationships between variables. They can give clues on the set of variables to search in order to find explanations for something happening in the network. This knowledge is highly useful in order to identify which variables should be used in the training phase. The domain experts have also the necessary knowledge to assess the obtained results and evaluate their significance.

These solutions correspond to a direct use of the domain experts' knowledge. In addition, it is also possible to use the domain expert knowledge indirectly through the creation of a qualitative model of the network. The qualitative model in the way used in this project (Camarinha & Martinelli, 1997) tries to mimic the way of thinking used by the domain experts. This qualitative model represents, in a simplified way, how the network's variables are interrelated and how they

influence each other. The qualitative model can be used as a tool to define which variables to use in the training phase and also to evaluate if the achieved results are consistent with the expected relationships between the variables. Figure 5 illustrates a partial qualitative model for a station.



Relations
$(G1 \mid G2) \& V5 \xrightarrow{-} N1$
$(G1 \mid G2) \& V5 \xrightarrow{+} P1$
$(G1 \mid G2) \& V5 \xrightarrow{+} Q1$
$G7 \& V4 \xrightarrow{+} P1$
$G7 \& V4 \xrightarrow{+} Q1$
$VC1 \& B1 \xrightarrow{-} HP1$
$VC2 \& B1 \xrightarrow{-} HP2$
$(VC1 \mid VC2) \& B1 \xrightarrow{+} P1$
$(VC1 \mid VC2) \& B1 \xrightarrow{+} Q1$
$B2 \xrightarrow{-} SA1$
$B2 \xrightarrow{+} P1$
$B2 \xrightarrow{+} Q1$
$VC3 \& B3 \xrightarrow{-} CF1$
$VC4 \& B3 \xrightarrow{-} CF2$
$(VC3 \mid VC4) \& B3 \xrightarrow{+} N2$
$Q1 \xrightarrow{+} N2$
$P1 \xrightarrow{+} Q1$

Figure 5 – A pumping station and part of its qualitative model

5.2 Users of the learning system.

The architecture of the learning system is composed by two separate modules in order to take advantage of the a priori knowledge of the domain experts, and also thinking in the further daily use of the system. The first one is used to acquire the knowledge, i.e., to select subsets of the historic data, feed them to the learning algorithms and extract knowledge from them. The second part is used to apply the acquired knowledge, i.e. to use it in the supervision of the water distribution.

The first component called "Knowledge Extractor", is one that requires a stronger interaction with the water distribution experts. These experts are, in general, people with a high education level and a strong knowledge of the network operation and physical distribution.

The second component, "Knowledge performer", which is the part that is going to be used more frequently, should be suitable to be used by the network operators. Sometimes these people do not have a high level of instruction. They acquired their know-how on the operation of the network with their past experience. The characteristics of these operators have to be taken into consideration during the development of this module. The tool should be as easy to use as possible and appealing only to the kind of knowledge they possess.

The knowledge extractor is going to be used sporadically. During the system installation phase, the domain expert is requested to define where and how to apply the available learning tasks. However, as the time goes on, the use of the knowledge extractor becomes restricted to refine the knowledge already acquired, to capture new operation characteristics or to learn about some new parts of the network.

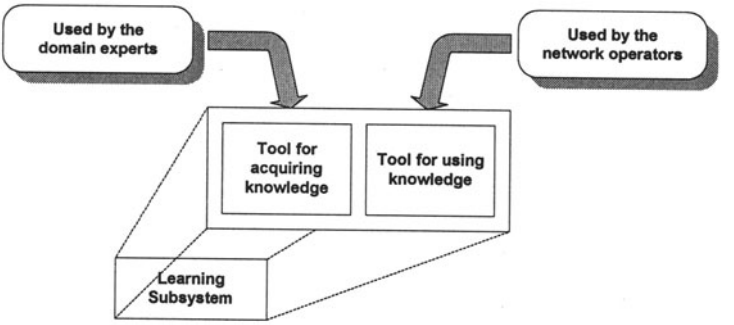


Figure 6 - Usage of the Learning Subsystem

The second component tends to be used continuously. The operators at least once a day, want to determine which is the predicted water demand in the network. Also each time an alarm occurs, the second module is activated to help solving the error situation.

5.3 Tools to assist the users

5.3.1 Synoptic diagram

The synoptic diagram is a graphical representation of the network highlighting the devices that request some kind of control and the measurement devices. The synoptic diagram is analogous to the one traditionally used by operators of the network.

In the knowledge extractor module this diagram is used to help the expert in the process of selecting the variables to be used. The use of the synoptic diagram represents an easy and convenient tool to enable the water domain experts identify and select the variables that he thinks would be useful to consider in the learning process.

5.3.2 Pre-defined pre-processings

Pre-processing applied to raw data before these data can feed the learning algorithms are determinant for the success of the learning phase. However there is a wide set of possible pre-processing procedures that could be used. The calculation of average values or derivatives, the transformation from flow values to consumed volumes, or the mapping of variables' values for a different domain, are only a few examples of them. How to cope in a flexible way with this amount of possibilities is a very difficult question to answer. It is not realistic to expect the users of this system to hand code these procedures in any programming language.

Probably a good approach to solve the problem is the development of a graphical tool that would enable the learning system's users to interactively develop pre-processing procedures based on some graphical primitive functions. However, this remains an open question.

The solution adopted in this phase of the project was the direct implementation of a minimal set of pre-processing procedures that covers the needs of a reasonable set of possible application areas. The set of pre-processing procedures was chosen based on the experience obtained with the preliminary experiments and with the interaction with the end users. Although this is not the best solution, it represents a good trade-off among ease of use, flexibility and the development time constraints. Further procedures can later be introduced in the system but this will require a programmer.

5.3.3 Qualitative models

The qualitative model tries to express, in a simplified way, the interaction between variables used by the operators and domain experts while evaluating network situations.

The qualitative model, in the way it is used in this project, can be easily extracted on a manual basis from the synoptic diagram. However it can also be automatically extracted from a more abstract representation of the network, like the MFM models (Lind, 1994; Larsson, 1996), if such models are available. A development in this direction is being pursued in cooperation with the Sebetia company.

The qualitative models can be used in two situations. When selecting variables, the qualitative model can suggest to the users which other variables have some relationship with a variable previously chosen. It is the responsibility of the system's user to accept or reject the suggestion based on his own knowledge. This mechanism increases the accuracy in selecting the set of variables to be used in the training phase.

The second usage of the qualitative model is to help in the assessment of the results of the learning algorithms. The mechanism is based on the evaluation of the rules extracted by the learning algorithms. If a generated rule represents relationships that are consistent with the relationships represented in the qualitative models, then the extracted rule is suggested to be correct. Again, the final judgment is the responsibility of the system's user.

5.4 Knowledge extractor

The knowledge extractor is the core of the learning subsystem. Its operation is based on 6 steps:

- Variables selection: The human expert selects the variables he/she wants to use in the learning process using the synoptic diagram. The qualitative model suggests some variables, but the human is responsible to accept or reject the suggestion.

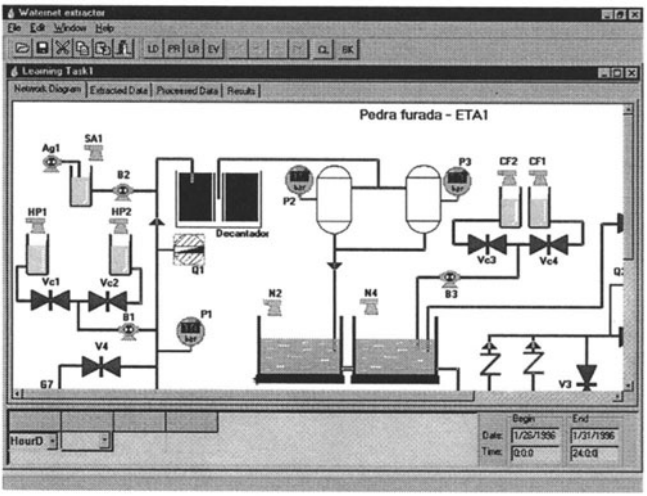


Figure 7 - Example of the use of the synoptic diagram in the knowledge extractor

- Data capture: The system extracts, from the DIMS, (Distribute Information Management System developed by the University of Amsterdam) the information it needs to perform the learning task.
- Pre-processing: Pre-defined pre-processing procedures are applied under human guidance in order to extract some high-level features that will ease the learning process and increase its accuracy.
- Data formatting: This phase closely follows the previous one. This phase also applies some pre-processing to the data. However the objective is to transform the data in a format that could be fed to the learning algorithms.
- Learning: For the learning phase, the user selects the execution of one of the learning algorithms available in the implemented catalog of learning techniques.
- Results assessment: The result assessment is performed in two different ways according to the learning task that the user is working with. If the user is working with the demand forecast, the results' assessment is done through the use of a graphic. For the error monitoring and alarm handling task, the assessment is performed using the qualitative model and the user judgement.

5.5 Using the Extracted Knowledge

Due to the different characteristics of the two selected learning tasks, the use of the extracted knowledge follows two different ways. One module to work with the water demand forecast and a knowledge based system to work with the error monitoring and alarm handling.

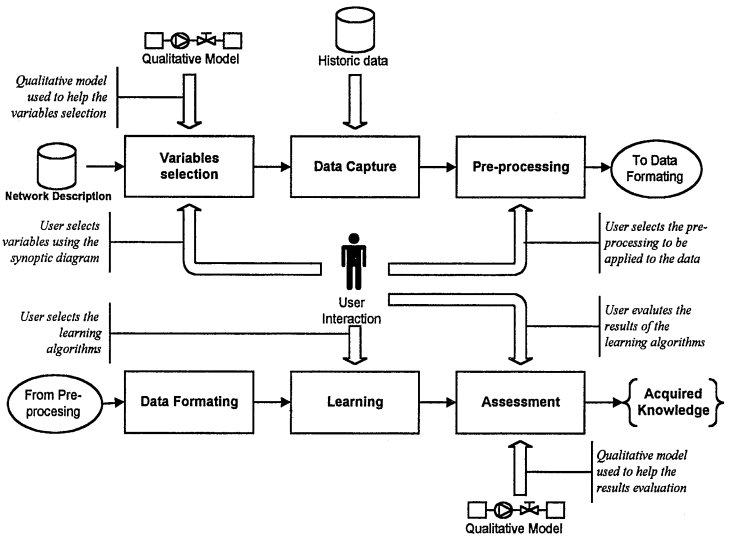


Figure 8 - Points of use of domain expert knowledge in the knowledge extractor

5.5.1 Water demand forecast

To work with the water demand forecast task, a module which takes the results of the knowledge extractor and uses them to predict the demands in various points of the network was designed.

The results of the knowledge extractor, regarding this application area, are represented in the form of neural network architectures. Thus, the work of the water demand forecast module is to take the neural network architectures and execute them considering the objectives of the operators and information about the current status of the network.

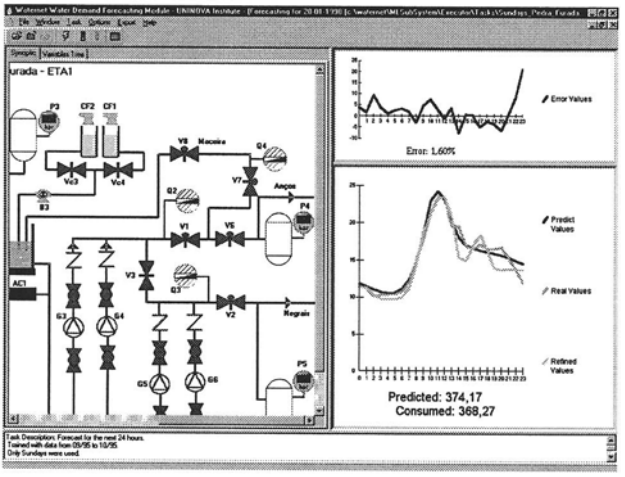


Figure 9 - Example of the water demand forecast module

To help the operators understand the process, a synoptic diagram is also used. The result of the execution is represented as a curve containing the predicted demand for the next hours. It is possible, in the same interface, to compare the predicted values with the real values collected directly from the network. It is also possible to continuously refine the predicted value using as new information the values being collected from the network. Figure 9 illustrates this process.

5.5.2 Error monitoring and alarm handling

Learned knowledge related to the error monitoring and alarm handling is represented in the form of rules resulting from inductive learning algorithms. The option for a rule-based representation instead of neural networks is motivated by the need to have the human experts' assessment. If, for instance, an alarm is detected, and according to the system status, one action has to be performed, the system operator has to be informed and advised about what is happening in the system. Neural networks, besides their goods aspects, are not good to tasks where explanation is required. Rule based systems are more appropriate for this purpose.

An example where inductive methods were applied to generate the rules is presented above. The following example takes the sensorial information from the station shown in the Figure 10.

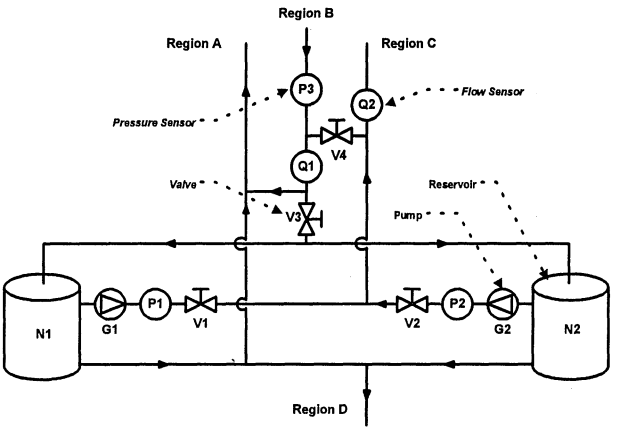


Figure 10 - Station diagram

The idea used in this learning example was to identify which operations were performed after an alarm firing, i.e. the operations performed in order to recover the system from the alarm situation. This experiment emphasises the need to transform the raw data into attributes of higher level. The data used in the learning process contained:

- System variables like pressures, flows, and device status;
- Time period between the moment in which a operation was performed and the present moment;

- Time period between the moment the last alarm was fired and the present moment;
- Operation performed in the present moment;
- Number of operations performed in the last 5 minutes;

Except for the first line, all the other lines represent attributes that are not explicitly gathered. They result from some kind of pre-processing applied over the raw data. The studied alarm refers to a problem in the valve V3. And the result represents the action that the operators execute. After presenting the examples to the C4.5rules (Quinlan, 1993) the result obtained was:

```
IF    OV3_MHigh_L5 < 3
THEN  Next_Action_Performed = M_High

IF    OV3_MHigh_L5 >= 3
THEN  Next_Action_Performed = M_Low
```

This means that when the mentioned alarm appears, the operator, in general, tries to open (M_High) repeatedly (3 times) the valve V3. If it still generates the alarm, then the operator closes the valve (M_Low) and tries to identify, in site, what caused the problem.

6. CONCLUSIONS

This paper described an application of machine learning techniques in the management of water distribution networks. The emphasis was put on the implementation of a practical methodology rather than on the development of new algorithms.

Preliminary experimental results shown that there is a large potential for application of learning techniques in a wide variety of sub-problems in the water distribution domain. A strong interaction with the domain experts and a step-by-step experimental approach seems to be an adequate method in order to get useful results.

This work is an ongoing work for which many open questions remain. Some of the next steps requiring further investigation include the use of qualitative models in the assessment of the extracted knowledge, and the support for more flexible pre-processing of the raw data, i.e. procedures to extract high level features.

Although a large number of potential learning tasks were identified, up to now only two of them, demand forecast and alarm handling, are being investigated. The other areas remain for further work.

ACKNOWLEDGEMENTS

This work is funded in part by the European Commission, via the Esprit Waternet project. The authors thank ESTEC and SMAS-Sintra for the supply of historic data and fruitful discussions.

The authors also thank the collaboration of the students Fernando Romano and Nuno Alves in the programming of the learning subsystem.

Fernando José Martinelli also thanks CNPq (Brazilian Council of Research and Development) for his scholarship.

7 REFERENCES

Afsarmanesh, H., Camarinha-Matos, L.M., Martinelli, F.J. (1997), Federated Knowledge Integration and Machine Learning in Water Distribution Networks, In *Re-engineering for Sustainable Industrial Production* (Camarinha-Matos, L.M., ed.), Chapman & Hall, London, 121-140, 1997.

Camarinha-Matos, L.M., Martinelli, F.J. (1997) Application of Machine Learning in Water Distribution Networks: An Initial Study, in *Proceedings of the Workshop on Machine Learning Application in the real world, Methodological Aspects and Implications* (Engels, R. et al., eds.), Nashville (TN-USA), 48-57, 1997.

Loucks, D. P, Costa, J. R. (1991) Computer-Aided Decision Support in Water Resource Planning and Management, in *Decision Support Systems: Water Resources Planning* (Loucks, D. P. and Costa, J. R., eds.), Springer-Verlag , Berlin(Germany), 3-41, 1991.

Larsson, J. E. (1996), Diagnosis based on explicit means-end models, in *Artificial Intelligence*,80, 29-93, 1996.

Lind, M. (1994), Modelling goals and functions of complex industrial plants, in *Applied Artificial Intelligence*, 8, 259-283, 1994.

Quevedo J. et al. (1987), A contribution to the interactive dispatching of water distribution system, in *International Symposium on AI, Expert Systems and Languages in Modelling and Simulation*, Barcelona, 41-46, 1987.

Quinlan, J.R. (1993), *C4.5: programs for machine learning*. Morgan Kaufmann Publishers inc., San Mateo(CA-USA), 1993.

8 BIOGRAPHY

Luis M. Camarinha-Matos is associate professor at the New University of Lisbon where he coordinates the Robotics and CIM group. He has been involved, both as researcher and has technical coordinator, on several international research projects in the areas of virtual enterprises, multiagent systems, intelligent manufacturing systems and machine learning. He has served in the Program Committee of many conferences and was one of the founders of the BASYS conferences series.

Fernando J. Martinelli received his M.Sc. in Electrical Engineering, emphasis on Automation, from Federal University of Espírito Santo (Brazil). He is currently getting his Ph.D. at the Department of Electrical Engineering of the New University of Lisbon. His research interests include machine learning, qualitative reasoning and supervision of distributed systems.