ALARM CORRELATION FOR COMPLEX TELECOMMUNICATION NETWORKS USING NEURAL NETWORKS AND SIGNAL PROCESSING

Emmanuel Marilly, Armen Aghasaryan, Stéphane Betgé-Brezet, Olivier Martinot, Gérard Delège

ALCATEL CIT
Route de Nozay, 91460 Marcoussis, France

Abstract
The aim of this paper is to investigate new tracks to tackle the open issues of alarm correlation for complex networks (e.g. IP/SDH or IP/WDM). The suggested approach consist of transforming, thanks to a neural network, the equipment alarm messages into a signal, and then to use signal-processing methods in order to extract the relevant information for diagnosis.

1. Introduction
Telecommunication networks are growing in size and complexity, which means that a bigger and bigger volume of notifications needs to be handled by the management systems. Most of this information is produced spontaneously by equipment and this message flow must be pre-processed to make effective management possible. Filters based on a per-notification basis fail to perform an adequate information pre-processing required by human operators or by management application software, which are not able to process such an amount of events. A pre-processing stage must decrease this information stream by suppressing superfluous notifications and/or by aggregating relevant ones.

In addition, the problem of expertise acquisition still remains the same: how to feed the filtering system? Where is the needed information available? And how to express this information?
The aim of this paper is to investigate a new alarm management approach with the finality of network malfunction recognition or fault recognition. The first part gives a brief overview on the state of the art of alarm correlation and diagnosis in telecommunication networks. The second part describes the basic procedure we suggest in this paper for alarm correlation.

2. State of the Art
Many different approaches exist in the literature and propose more or less complex intelligent filtering: one can use some efficient rule-based languages, and/or object-based techniques like ECXpert (Event Correlation eXpert) which builds alarm correlation trees according to some handwritten rules [NY95]. Generally alarm correlation techniques can be classified within two main groups.

- Symptom-based approaches which rely on the visible/observable phenomena or signatures, e.g. alarm patterns, to describe each possible fault situation. Expert systems relying on a set of rules belonging to this first category, as well as chronicle-based approaches which consider quantitative temporal relations between the observed events [DO96].

- Model-based approaches which rely on a more systematic study of the supervised system in order to model its structure and its behaviour. Then, these models can be used to explain the on-line alarm observations by deducing the underlying faults, and fault propagations. An example of such an approach can be found in [AG98], where a randomised Petri Net model is used by a dynamic
programming algorithm for fault diagnosis. The models can also be used to build a simulator for generating alarm scenarios [LK97] that afterwards could be used on-line by symptom-based approaches to identify the underlying fault.

Today, the symptom-based approach is the most commonly used in the supervision of existing telecommunication networks. It is this approach that we consider in this paper.

One of the main difficulties of this approach is the acquisition of the expertise-based knowledge. Updating this expertise according to the evolution of the system is clearly also a problem which has to be taken into account when components are subject to frequent changes (topological or functional). When the field of training changes, significant maintenance is necessary with adaptation and modification of the rules. The evolution of these systems to a non-expert person is not obvious. The expert systems do not support the use of new data (unknown data) or the change of initial data. If an event is not received (data missing), the rule is not applicable any more. The rules are not robust. These methods do not learn anything of their experiments and adapt with difficulty to the evolution of the network. These techniques are not good dealing with the uncertain and they have difficulties to analyse a significant number of data not correlated, ambiguous and incomplete...

Within the same approach, Gardner [GA97] has proposed a solution using artificial neural networks for alarm correlation. The main disadvantage of its solution is that the time relation between alarms is not taken into account.

3. Basic Solution Description

The method that we propose is a hybrid method. The neural networks and signal processing techniques are combined to propose a new solution of alarm management and diagnosis. The objectives are to use the advantages of both technology (i.e. robustness, fault tolerance, learning etc…) and to take into account the time relation between alarms i.e. process sequences of alarms.

The aim of the first section of this paragraph is to define the neural network terminology for non-specialists [RM 86]. The second section describes the global process of the proposed method.

3.1 The neural network

Connectionist models are also called neural networks. However most current neural network architectures do not try to closely imitate the biological model but rather can be regarded simply as a class of parallel algorithms. A network consists of units and directed/weighted links (connections) between them [HE90]. In analogy to activation passing in biological neurons, each unit receives a net input that is computed from the weighted outputs of prior units with connections to leading to this unit. The knowledge is usually distributed throughout the network and is stored in the structure of the topology and the weights of the links. The network is organized by training methods (adjustment of the weights of the links to get the desired system behaviour).

Classical logic in AI systems is replaced by vague conclusions and associative recall. This is an advantage in all situations where no clear set of rules can be given. The inherent fault tolerance of connectionist models is another advantage. Furthermore, neural networks can be made tolerant against noise in the input. One of the major advantages of neural network is their ability to generalize. This means that a trained network could classify data from the same class as the learning data that it has never seen before.

3.2 Global process of the solution

This section describes the different processing steps of our hybrid solution. Five steps have been identified that could be regrouped in three families: the pre-processing (i.e. data analysis), the processing (the algorithms), and the validation (validate the diagnosis). These three groups correspond to the on-line/off-line process (the processing is on-line, the validation and the data analysis are off-line). Figure 1 gives a diagram of the basic idea. The different steps with their processing are the following:

- The step 1 concerns the definition of the alarm log syntaxes. The alarm logs (generated by the network and currently ASCII words) are coded in binary data according to a specific organization (syntax of coding). An example of mapping between the alarms and their binary words is described in the next paragraph. The redundant alarms are erased. The definition of the mapping between the alarms and their binary word is performed off-line. After this definition is done, the mapping is processed on-line.

- The step 2 concerns the analysis of the different binary words provided by previous operation to identify alarm classes. The aim of this analysis is to create the different classes used for the training set of the neural network. For each binary word defined previously an identification class is defined (several binary words can be associated to the same class). To determine these classes, techniques such as principal component analysis can be used. This process is done off-line.
The step 3 concerns the neural network processing. The neural network classifies the different binary words corresponding to the alarm logs into the classes defined previously. A multi-layer feedforward neural network [RM 86] is used to process this transformation. These neural networks have an ability to learn what is carried out by a reinforcement of connections between the neurons. The learning is performed off-line using examples. Once the learning finished, the classification is made on-line.

The step 4 concerns the signal processing. The signal is provided by the output of the neural network. It corresponds to the class evolution with the time (see paragraph 4 for more information). The signals contain all the pertinent information necessary to do the diagnosis. Advanced signal processing methods can be used to extract this pertinent information. The pertinent information will not only concern unique alarms but also sequence of alarms (richer in information, more robust). According the information requested, several signal processing techniques can be used. Three groups of different analysis methods can be carried out:

- The use of Fourier analysis techniques changes for instance a windowed signal into frequency spectra. This is a frequency analysis method and the extracted information concerns a frequency in the signal. A specific frequency can correspond to a specific malfunction.
- The use of time/frequency analysis techniques (for instance Wigner-Ville distribution [CO89]) produces a bilinear energy distribution of the signal. The results obtained through this distribution will be interpretable in term of times and frequencies according to a linear scale.
- The use of time-scale analysis techniques (for instance the wavelet transform) changes the signal into coefficients which can be used for analysis and interpretation. This analysis uses the resolution properties (different level of resolution for the analysis).

The final diagnosis will be carried out either directly starting from the observation of the analysis tools results or with the help of another tool. This tool can use neural networks, vector quantization (or other) to classify and produce the result. This step has an off-line part (identification of the pertinent information) and an online part (the recognition).

The step 5 concerns the validation of this method on real use cases. The verification of several aspects (validation of the detected malfunction, reproducibility, robustness etc...) of our method must be done. If the validation provides bad results, step 2 must be modified. The validation step is only off-line.

Now that we have explained the methodology, we apply it on a concrete example in the next paragraph.

4. Example

To illustrate the example, the method described above is applied on a specific failure of an SDH network. The failure introduced is a cut in an optic fiber between two SDH equipments. The goal of this failure is to produce several hundred of alarms.

For each generated alarm, the operator receives a message with the following format:

```
09/09/1999 09:51:24 1651SM24 3x34M TRIB 4   URG
18ALE AU15A300M01  CARD MISMATCH DEBUT
```

![Figure 1: Basic Description of the proposed solution](image-url)
These alarm messages are generated according to the Nectas format (Alcatel proprietary). These alarms are modified and coded (mapping of step 1) so that a neural network can use them. The coding is processed in the following manner:

- The date and time are eliminated because the network of neuron is a function of the time.
- With each type of alarm and each logical entity, an input neuron is associated.
  
  Example: The entity 3x34 M TRIB 4 is associated with input neuron number 8 and the alarm CARD MISMATCH is associated with input neuron number 39. When the alarm of this entity becomes active neurons 8 and 30 pass to 1 (0 when inactive).
- A neuron codes the perceived severity.

Once the alarm logs are mapped into binary words, an analysis is done (step 2) to associate these ‘words’ with a class. The mapped alarms, associated with their classes, constitute the pattern training of the neural network. The properties of the neural network must be defined before using it. The selected network is a multi-layer feedforward neural network and the supervised learning algorithm is the scaled conjugate gradient [Mol 93]. The topology of the neural network is the following: one input layer, two hidden layers and one output layer. Taking into account the number of logical entities, alarms and perceived severity for this experiment, the neural network has 64 neurons in the input layer, 129 neurons in each hidden layer and 5 neurons in the input layer. The training pattern has 79 elements.

An optical fibre cut between two SDH equipments produces several hundreds of alarm messages. The use of a neural network (step 3) transforms these messages for each equipment into a signal (see figure 2). Figure 2 shows the alarm classes evolution with time (the alarm logs are classified by the neural network and the output of the neural network is drawing throw the time). With this alarm behaviour the visualization is easier and the signal processing tools enable interpretation and diagnosis.

To contribute to the interpretation of this behaviour (see figure 2) it is then necessary to do a signal processing analysis (step 4). As an example, figure 3 shows the result of a signal processing analysis by a time frequency analysis.

In this case, the relevant information contained in the signal corresponds to peaks in the time-frequency representation. The localization of the peaks maxima can be extracted. Each peak gives a vector (time, frequency and energy localization). These vectors correspond to the signatures of the breakdowns and are used to perform the diagnosis.
If we use a more basic signal processing technique such as a power spectral density (see figure 3, the energy spectral density) the localisation of the spectrum peaks can be used to identify the malfunction.

Finally, each identified solution must be validated. This validation consists of an experimental verification. For the same failure reproduced several times, the malfunction must be well identified. This validation step (step 5) must be done by a network expert.

6. Advantages / disadvantages of the new solution

Our solution has not the objective to replace the existing solutions but provides an indication of how to solve some open issues of the current methods.

The main advantages are:

- Doing the diagnostic not only on classes of alarm but also on sequences of alarm. This means that the time information is taken into consideration. Performing the diagnosis on alarms associated with the time is more complex but provide better results for the final diagnosis.
- Visualisation. Alarm logs are not seen as messages but as signals. An interpretation of a graphical representation is usually easier than an interpretation of ASCII alarm logs. The behaviour obtained by the neural network allows to have a complete visualization of all active alarms at any moment.
- Robustness. The loss or the alteration of some messages does not involve a significant modification of the signal and does not deteriorate the diagnostic process. Our process is robust because we take into account a set of data. This robustness is due to the properties of the neural network and signal processing techniques.
- Generalisation capabilities. All the neural networks have the generalisation capability (see paragraph 3.1)

This solution obviously has some disadvantages as well:

- The identification of malfunctions in signal and alarm logs needs experts.
- Detection of a new failure or malfunction needs more development than a classic rule engine.

7. Conclusion

We have presented an approach for network alarm diagnosis that combines neural networks and signal processing techniques. They are already used separately, but a hybrid solution strengthens the advantages of these two techniques. However, this robust solution is not easy to implement and requires experts. There is a compromise between development and robustness.

Finally, although the experiment was carried out for an SDH network, we can reasonably suppose that the approach can be extended successfully to other networks (such as WDM, WDM/IP...). The aim of this solution is not to take the place of the existing solution but to provide a complementary way to do the diagnosis.

8. References