# Neural Networks for QoS Network Management

Rafael del-Hoyo-Alonso[1], Pilar Fernández-de-Alarcón[1],
Juan-José Navamuel-Castillo[1], Nicolás J. Medrano-Marqués[2],
Bonifacio Martin-del-Brio[2], Julián Fernández-Navajas[2], and David Abadía-Gallego[1]

[1] Instituto Tecnológico de Aragón, C/ María Luna nº6,
50018 Zaragoza, Spain
{rdelhoyo,pfernandez,jnavamuel,dabadia}@ita.es
[2] Departamento de Electrónica y Comunicaciones, University of Zaragoza,
C/ María Luna 1, 50018 Zaragoza, Spain
{nmedrano,bmb,navajas}@unizar.es

**Abstract.** In this paper we explore the interest of computational intelligence tools in the management of heterogeneous communication networks, specifically to predict congestion, failures and other anomalies in the network that may eventually lead to degradation of the quality of offered services. We show two different applications based on neural and neurofuzzy systems for Quality of Service (QoS) management in next generation networks for V2oIP services. The two examples explained in this paper attempt to predict the communication network resources for new incoming calls, and visualizing by means of self-organizing maps the QoS of a communication network.

**Keywords:** Intelligent network, Call Access Control, IP Networks, QoS, Neural Networks, Neuro-Fuzzy Systems, EFUNN, Self-Organizing Maps.

## 1 Introduction

In a close future, Internet will become the unique platform for voice and video services. However, the quality of multimedia services offered on the Internet depends of congestions, failures and other anomalies in the network. Today, it is usual to find these problems since the Internet is a mixture of heterogeneous networks, and the quality is affected by very complex high non linear problems. Furthermore, the success of new services based on the Internet depends on how internet service providers guarantee the quality of service (QoS) [1], which is defined subjectively as 'the collective effect of service performance which determines the degree of satisfaction of a user of the service'.

In order to achieve quality of service requirements it is necessary to involve resource allocation. On the other hand, network efficiency crucially depends on the degree of resource overbooking inside the network. A key problem in concurrently achieving both goals is caused by the fluctuation over multiple time-scales of the traffic load emitted by multimedia applications, and high dimensionality of network statistics used for management.

The main idea of this paper is the use of techniques from the computational intelligence field in order to deal with high dimensionality information, and to

F. Sandoval et al. (Eds.): IWANN 2007, LNCS 4507, pp. 887–894, 2007.
© Springer-Verlag Berlin Heidelberg 2007

improve network behavior based on the prediction of the QoS. These techniques have been implemented in a new network element, called Prediction Resource Manager, in the framework of the Quar2 European project [2]. One feature of this element is its capability to observe in real time the network ‘quality’ by using self-organizing maps for a straightforward representation of available resources in the network. Also the proposed element will use neural networks to predict the available resources of the network. A goal is the prediction of resource availability, with sufficient accuracy, in order to allow admission control of new multimedia flows in such a way that non QoS calls will be rejected, and the quality degradation of current multimedia flows will be avoided.

The presented results have been developed within the Quar2 European project [2], framework for this paper, and whose objective is to offer full interactive multimedia services guaranteeing the long desired, yet hard to achieve, toll-quality voice and video service over heterogeneous IP (Internet Protocol) networks (V2oIP) and over heterogeneous network environments, focusing on the access network.

The paper is organized as follows. In Sect 2 we briefly describe the problem and different technologies used in our research. In Sect. 3 we show our approach to de problem, and the neural architectures and tools developed. In Sect. 4 we discuss our results in QoS visualization and prediction. Finally some conclusions are provided.

## 2 QoS and Neural Technologies

### 2.1 Quality of Service and Traffic Prediction in Heterogeneous Networks

QoS measuring and resource prediction is one objective of this research, which is a classical problem in the area of distributed multimedia systems [3]. There are some proposals in the literature for traffic prediction [4, 5], being the goal to forecast future traffic variations as accurately as possible to predict the user network behavior. On one hand, a large prediction interval is needed to provide sufficient time for control actions, and to compensate the unavoidable delays caused by traffic measurement (sampling smoothing) and traffic prediction. On the other hand, a small prediction error is desirable for the admission of future calls in order to avoid quality degradation for incoming calls (Quar2 project objective).

The prediction algorithms used in our research are based in soft-computing techniques, due to non-linear operation, flexibility and learning from example capabilities. The input information for prediction is known communication network statistics (transport domain) from distributed network elements, and the perceived quality (service domain) from the running multimedia flows. A good approach to evaluate the network resources available in a best-effort network is to achieve the perceived quality of the running calls (video and audio flows); thus, if the QoS of the current calls in the network decreases, the network is in a congestion situation.

### 2.2 Neural Network Approach

Learning is one of the capabilities that make artificial neural networks and neuro-fuzzy systems a favorable approach for time traffic prediction. Supervised neural

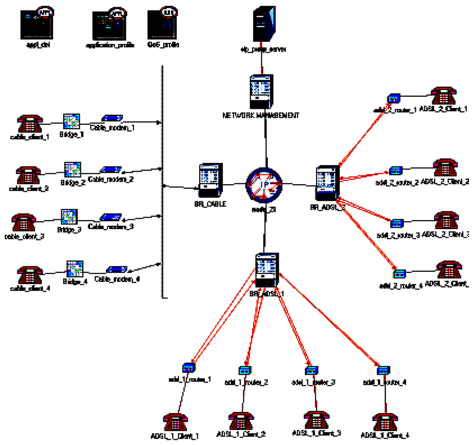


**Fig. 1.** Simple scenario A, with two ADSL access networks, and one cable access network

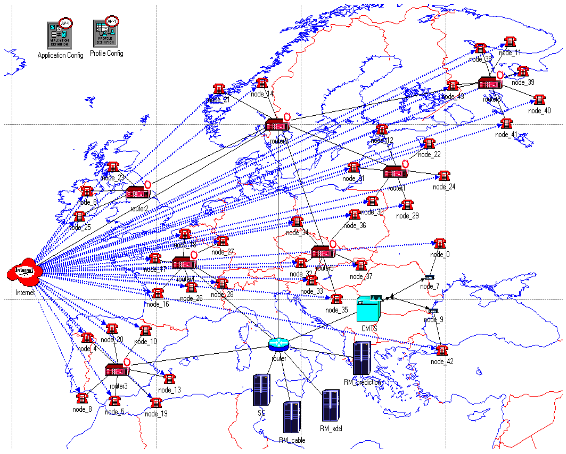


**Fig. 2.** Final Scenario B, with 48 clients, several access networks, and a simulated backbone

models, as Multilayer Perceptron (MLP), and Neuro-Fuzzy [6], as Evolving Fuzzy Neural Network (EFUNN) [7], have demonstrated their capability for solving supervised time prediction problems in many applications. We have applied EFUNN in our work because of the good results obtained in time-series prediction, and because it will also make possible the extraction of underlying classification rules, for future network troubles resolution. Traditional (linear) approaches, like auto regressive methods (AR, ARIMA) provide reasonable results in time prediction [3]. However, in a real situation, the required information for a time prediction task frequently is not completely available, or is not acquired at the same time. Incomplete data, great heterogeneity of information used, or high dimensionality of input information, are common problems in real data. Thus, neural and neuro-fuzzy systems could be a better approach for network management than traditional approaches.

On the other hand, we use Self-Organizing Maps (SOM) [8] for the visualization of the status of the communication network. We take advantage of the SOM capability of reducing the data dimensionality while retaining as much information as possible: As it is well known the SOM can be viewed as a non linear extension of principal component analysis (PCA). It replaces the linear subspace of PCA by a nonlinear manifold that can represent even very non-linear data distribution.

We will use the SOM to cluster the network statistics, and to visualize in a 2D diagram (map) in real time when the network has overbooked resources, or when the network has QoS problems.

## 3 Approach and Tools

The solution presented here aims to offer real-time operation among users located in different access network environments interconnected by an Internet Protocol (IP) backbone. The Prediction Resource Manager is the element which estimates the behavior of the communication network. Its main purpose is to predict from network statistics the QoS of the backbone, and visualize network QoS status of current calls.

The proposed approach uses IP router statistics as input (IP packets discarded, bytes/packet forwarded, packets interfaces in and outs). Predictions based on MLP and EFUNN, and the visual information provided by the SOM, can be viewed as aids for the ulterior human treatment and decision.

Our basic scenario consists of several IP phone clients calling among them. All calls have to be admitted by the service controller (main entity in an IP telephony service domain), which asks for resources to the Prediction Resource Manager (transport domain). Generally, a pending client will be admitted when sufficient system resources are available; that means in this context that the Mean Opinion Score (MOS), a QoS measurement, of incoming calls will be fine. Thus, for the admitted clients, as well as for a pending client session, the future consumption MOS rates have to be estimated in advance.

First, the SOM visualization allows us to find a map region where the admission is possible, and to visualize the network status. The objective is to represent in real time the status of the communication network: the SOM provides a straightforward visualization of the current network resources, and allows determining when the network can admit new incoming calls or when it is not possible. Then, the second neural network for traffic prediction forecasts the quality of the new clients for call admission of new incoming calls in order to make an automatic decision system.

The Intelligent Resource Manager element uses the information of the network status as input and, by using the different strategies presented above, allows the system to find a final decision in order to accept or not a new call. MLP and EFUNN were implemented by using a library called Waikato Environment for Knowledge Analysis (WEKA). This is a comprehensive suite of Java class libraries that implements many state-of-the-art machine learning and data mining algorithms. We also use the DataBionic ESOM tools library for the SOM, which offers several SOM graphic interfaces (U-Matrix, P-Matrix, Component Planes, SDH...) [9].

All data network statistics are treated as a set of data items. A dataset is a collection of network statistics, where each pattern consists of a number of attributes (example, router IP packets discarded), the last attribute corresponds with the QoS evaluated from currents calls (this value is a global indicator for the network quality). It can be qualitative (one of a predefined list of values example, quality gold silver or platinum), or numeric (a real or integer number).

OPNET Modeler, a telecommunication simulator platform (leader in network simulation tools), has been used in order to generate IP complex scenarios. We use OPNET for simulating a full heterogeneous architecture in order to evaluate QoS mechanisms, timing prediction algorithms, and network visualization tools.

We have developed an interface between OPNET Modeler and an external application for the Intelligent Resource Manager. This new feature allows OPNET to incorporate a new tool for statistics visualization, making possible to interoperate with the network to create a call control for simulated applications. The OPNET link and the software designed allow exchanging easily the simulator for a real network link.

## 4 Results

Following this approach, and by means of the developed tools described in Sect. 3, two studies have been carried out. First, a QoS map for visualization and prediction in a simple scenario (Fig. 1) has been developed. This scenario consists of 12 clients with restricted IP links, and no application or background traffic have been defined, so only VoIP traffic is simulated. The number of calls has been increased during simulation; two possibilities were defined for new incoming calls, with QoS or without QoS. The Prediction Resource Manager predicts and visualizes the network status in order to allow or reject new incoming calls.

The second and more complex scenario (Fig. 2) contains 48 clients making calls among each other using a fixed pattern. In parallel, background traffic was defined in order to simulate non-QoS application network traffic. This second scenario simulated access network, Cable and ADSL, with different network congestions. This complex scenario simulates Quar2 architecture with all its elements.

### 4.1 Visualization

In order to generate a tool for network QoS visualization, each statistics group for the network elements in a time period was associated with the MOS average of the current calls, and then associated to non-QoS class or Gold QoS class.

We use as criterion for the SOM development the minimum square error (MSE) for the training data. The best MSE was achieved with an 82x50 neurons map (related to the two first principal data components), after 200 training cycles.

In the SOM map obtained for the scenario A (image not included due to space restrictions) two quite separated regions appeared. It is possible to discriminate how non-QoS and QoS regions are defined, so we can classify when the network has QoS or not for call admission. These maps are a conceptual description of the network, and a good representation of QoS status.

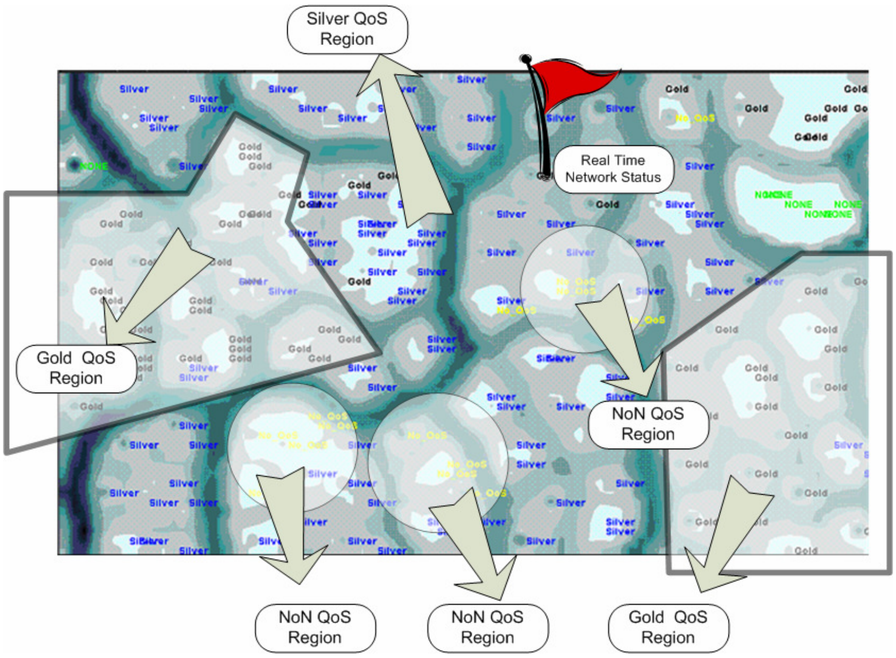


**Fig. 3.** SOM Two-Match Background visualization for scenario B. The map represents different QoS regions: high quality (gold), medium quality (silver) and non QoS.

In the second scenario, the statistics and network behavior are more complex than in the previous one. In this case three different classes are pre-defined: GOLD, SILVER and NON_QoS. In the SOM visualization (Fig 3), a number of regions are identified, and it is possible to distinguish how the NON_QoS, SILVER, and GOLD quality regions (standard discretization for quality) are defined; hence, it is possible to classify when the network has QoS or not, and the predicted Service Level Agreement (SLA). In that case, the regions are not defined as clear as in the previous scenario due to the complexity of the network and the SLA definition (gold, silver...). However, these maps can be also interpreted as a conceptual description of our network, and are a good representation of its QoS status..

The rough lines that separate the sub-regions in the map of Fig. 3 represent an interpretation of the distance between the different neurons of the map, i.e., the distance between clusters (called Two-Match visualization). Furthermore, different legends for a better interpretation of the map have been introduced. The different labels correspond to points of the map labeled by using a historic dataset.

Finally, the areas sketched represent regions of the map with QoS and without QoS. Finally, a flag representing in real time the state of the network (Fig 3) was introduced during the simulation (it moves along the map).

### 4.2 Prediction

Several regression methods were finally tested for prediction; EFUNN [7] and MLP were used to predict the average quality of the communication network. The MLP and

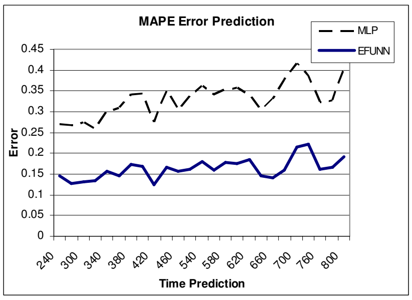


**Fig. 4.** Mean Absolute Percentage Error (MAPE) prediction error calculated for Multilayer Perceptron and EFUNN. The x-axe represents time in advance future prediction (from 240 to 800 seconds).

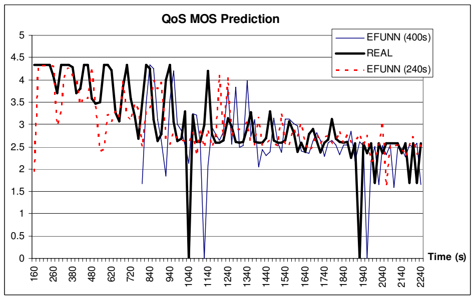


**Fig. 5.** Prediction obtained using EFUNN algorithm. Two predictions are represented, plotting future QoS prediction error from 160 to 2240 seconds. EFUNN is used for prediction 240 and 400 seconds in advance.

EFUNN models were trained from the historic data and evaluated with new incoming time network statistic instances. Once new QoS for the last instance is measured, this value is introduced for training. The evaluation of the prediction algorithm was always done with new temporal instances, and different prediction experiments were executed predicting from 240 to 800 seconds in advance.

For reducing the neural network size and avoiding overfitting, input data for both neural networks were reduced from 115 inputs to 10 by using Principal Component Analysis (PCA). The EFUNN algorithm [7] was configured with 3 function members and 100 seconds of node age. On the other hand, the learning algorithm chosen for the MLP was the gradient descent algorithm as implemented in the WEKA library. In order to achieve the best results, 100 network configurations were tested for MLP. Finally, the selected architecture uses two hidden layers, with 10 and 5 neurons, respectively.

The EFUNN algorithm achieved a prediction error lesser than a 20%, which allows generate a call access control. The EFUNN turns out to provide better performance than the MLP algorithm; we think that the reason is that EFUNN is more flexible than MLP, adapting better to random and forced condition changes. When new calls or packets are generated, all algorithms fail (provide large errors), but EFUNN adapts faster to the new conditions.

On the other hand, if the antecedents of network status make possible predicting congestion, MLP can do a better generalization. Furthermore, when plotting how the error increases versus time prediction, this approach results to be good enough for a short time prediction. Long time prediction will be studied in future research.

Finally, the EFUNN algorithm not only provides a better solution, but it also makes possible knowledge extraction. Thus, by the extraction of the underlying rules, appropriate fuzzy inference methods could then be applied on those extracted rules making possible to use a fuzzy rule-based system for the prediction task. This is a current topic of our research. In the Quar2 project, real results from the testbed have been studied with similar results obtained.

## 5 Conclusions

Real communication networks offer large amounts of information which must be analyzed and interpreted by experts and management applications for proactive and corrective management. Current algorithms analyze real time statistics for predicting events. Research in this field is carried out for discovering patterns of sequences, and time series and statistical analysis. This paper has presented several prediction methods based on neural network algorithms (EFUNN, MLP); they seem to be a valuable tool for predicting short-term as well as midterm target events when patterns exist. Besides, SOM visualization tools have turned out to be useful as a conceptual description for communication networks, and a good representation of its QoS status.

## Acknowledgments

This work has been supported in part by the Quar2 European project, and the Grupo de Ingeniería Avanzada (GIA) of the Technological Institute of Aragón.

## References

1. Wang, Z.: Internet QoS: Architectures and Mechanisms for Quality of Service. Morgan Kaufmann Publishers, San Francisco (2001)
2. Quar2 project, http://projects.celtic-initiative.org/QUAR2/description.htm
3. Sang, A., Li, S.: A predictability analysis of network traffic. In: Proc. IEEE INFOCOM (2000)
4. Morato, D., Aracil, J., Diez, L.A.: On linear prediction of Internet traffic for packet and burst switching network. In: Proc. 10th Int.Conf. on Computer Comms Networks, pp. 138-143 (2001)
5. Papagiannaki, D., Taft, N., Zhang, Z.-L., Diot, C.: Long-Term Forecasting of Internet Backbone Traffic: Observations and Initial Models. In: Proc. IEEE INFOCOM (2003)
6. Abraham, A.: Neuro Fuzzy Systems: State-of-the-Art Modeling Techniques. In: IWANN 2001. LNCS, vol. 2084, pp. 269–276. Springer, Heidelberg (2001)
7. Kasabov, N.: Evolving Fuzzy Neural Network for Supervised/Unsupervised On-line, Knowledge-based Learning. IEEE Trans. on Man, Machine and Cybernetics (2001)
8. Kohonen, T.: Self-Organizing Maps, 3rd edn. Springer, Berlin (2001)
9. Ultsch, A., Moutarde, F.: U*F Clustering: a new performace cluster-mining method based on segmentation of SOM. In: Proc. Workshop on Self-Organizing Maps, Paris, pp. 25-32 (2005)