A Scheme for Adaptive Self-Diagnosis of QoS Degradation in Future Networks

Panagiotis Spapis, Panagiotis Theodoropoulos,
George Katsikas, Nancy Alonistioti
Department of Informatics and Telecommunications
National and Kapodistrian University of Athens
Athens, Greece
{pspapis, p.theodoropoulos, katsikas, nancy}@di.uoa.gr

Stylianos Georgoulas
Centre for Communication Systems Research, Faculty of
Engineering and Physical Sciences, University of Surrey,
Guildford, Surrey, GU2 7XH, United Kingdom
s.georgoulas@surrey.ac.uk

Abstract—The capability of a network to identify problematic situations, named self-diagnosis, enables it to react promptly and autonomously once an event or error has been identified. The use of service information in this process enables it to identify more composite problems and to act more targeted in order to solve complex errors. This paper proposes a novel fuzzy logic-based self-diagnosis mechanism for identifying Quality of Service (QoS) degradation events. Furthermore, we introduce a framework for the adaptation of the self-diagnosis scheme, which enables the network elements to evolve the way they interpret the context information. The adaptation scheme is based on the statistical analysis of the measurements and reacts accordingly without requiring any external human intervention. The adaptive self-diagnosis scheme has been evaluated through simulations in order to showcase the benefits from its application in IP networks for the VoIP service. The simulation results show that the adaptive self-diagnosis scheme performs very well compared to existing solutions, increases significantly the event detection rate and, as a result, the capability of controlling the QoS on top of the involved network elements.

Keywords—Adaptive Self-Diagnosis, Statistical Analysis, QoS

I. Introduction

The acute proliferation of networking devices and advanced services, coupled with the demand for "anytime" and "anywhere" communications has increased the complexity of the contemporary networks. Thus, new sophisticated mechanisms for decision making procedures, using inputs stemming from several sources are required. A major requirement for the network elements as described and reflected in several recent research and standardization attempts is their ability to identify problematic situations during the operational runtime; such ability is named self-diagnosis.

Taking into consideration the complexity and the importance of the diagnosis problem, several approaches have been proposed in the literature. In [1] and [2] model based techniques are proposed, using logic backward reasoning to automate fault detection and localization based on Structural and Behavioral Model (SBM), by comparing observed network measurements with expected network behaviors and by tracing back causality structures. In [3] a mechanism for identifying users that misbehave in the network is presented

for access network elements. The mechanism uses network metrics (maximum, consecutive and actual back-off values) and based on their statistical analysis proceeds in detection of the users (and consequently service flows) that misbehave. The aforementioned mechanisms identify the problematic behavior of the nodes using network data and by correlating them with certain behaviors. In this paper we propose a fuzzy logic-based adaptive self-diagnosis mechanism for QoS degradation in future networks which uses both network and service layer information in order to make more accurate decisions. This work builds on our previous work in [4], where we have proposed a learning technique which adapts a fuzzy logic decision making mechanism for access network elements. We propose the introduction of a sophisticated adaptation scheme in order to enable the network elements to evolve the way they perceive their environment. The main differentiation regarding the adaptation mechanism compared to [4] is that instead of using simple clustering schemes, we propose the introduction of the statistical analysis of the measurements' dataset so as to capture its special characteristics and proceed in more sophisticated adaptations.

The rest of this paper is organized as follows: Section II gives background information related to the used statistical analysis and describes the proposed self diagnosis scheme with the adaptive framework. The performance analysis that showcases the applicability of the solution and the benefits from its incorporation in future networks is presented in Section III. The paper concludes in Section IV, where the main findings are summarized and directions for future work are given.

II. Adaptive Self-Diagnosis

A. Statistical Analysis Background

The statistical analysis and specifically the extraction of the distribution functions of a dataset provide crucial information regarding the special characteristics of the dataset. Many distribution functions are well described in the literature (e.g., Gaussian, Pareto, Log-normal, etc.). From the aforementioned distributions, the Gaussian is a very well-known one and is used widely in the literature for the description of various datasets [5]. The reasons for such acceptance are related to its nice behavior and formal

mathematical description. However, it is not ensured that all the information that could be derived from the dataset is being captured by a distribution. A way to tackle this problem is to use the Kernel Density Estimator, which is a non-parametric way to calculate the distribution function of a given dataset [6]. By using this technique the resulting distribution curve fits better on the histogram of the corresponding data without making any assumptions about the data. The Kernel Density Estimator of a sample $(x_1, x_2, \ldots, x_n)$ with an unknown distribution is described as follows:

$$\hat{f}(x) = \frac{1}{n \cdot h} \sum_{i=1}^{n} K(\frac{x - x_i}{h}) \qquad (1)$$

Where $K(\bullet)$ is the kernel and $h > 0$ is a smoothing parameter.

B. Self-Diagnosis in IP Networks using Service Information

Given the fact that Fuzzy logic is an ideal tool for developing diagnosis mechanisms, when dealing with contradicting or imprecise data, we propose the introduction of a self-diagnosis mechanism which will allow the management system, using inputs from network elements (network and service data) to proceed in identification of QoS degradation events in IP networks [7]. More specifically, we propose the introduction of a scheme that will exploit service information and by combining it with network measurements will identify whether a given service flow experiences a problematic situation or not.

Each service has its specific characteristics and requirements, and consequently, the network management system should monitor each service with its key performance indicators in order to identify problematic cases. Such service information can be obtained from the IP Multimedia Subsystem (IMS) servers (i.e. Proxy, Interrogating and Serving Call Session Control Functions – P-CSCF, I-CSCF and S-CSCF respectively). Each terminal is being registered to the S-CSCF through the P-CSCF (that acts as end point of the IMS network) and the I-CSCF (for communication with the S-CSCF). Similar procedure is being followed for the initiation of a service among IMS terminals – thus the IMS gathers information for the service that each user accesses; the afore-described information about which service is being accessed by every user is being kept in dedicated application servers. Each router in the path monitors the identified key performance indicators for the user-service pair and reports this information to the network management system. Then the network management system (which is in charge of carrying out the QoS diagnosis process) will identify the flows where a specific remedial action should take place in order to tackle a low QoS event.

For the case study under consideration and for the VoIP service, delay, jitter and packet loss are identified as playing significant role in QoS degradation, thus the event identifier evaluates the aforementioned inputs for every active session; for other services (e.g., highly reliable communications, IPTV etc.) other monitoring inputs could be used. The membership functions (MFs) of the identified parameters describe the input and output parameters. More specifically, the membership

functions indicate the values of each parameter, the range of each value and the magnitude of their participation. The "Delay", the "Jitter" and the "Packet Loss" per flow comprise the input vector, whereas the output is the "QoS degradation" [8][9][10]. The MFs' shapes chosen for the representation of the degree of certainty are trapezoidal, mainly for simplicity reasons and to so as to exploit the certainty areas for such inputs. For the QoS level the Gaussian membership functions are being used. The idea behind such adoption is mainly based on the smooth (i.e. the QoS should be related to the inputs in a smooth manner without non-linear alterations) and non-zero (the decision maker needs to conclude to a decision based on all inputs' range) nature at all points; for simplicity reasons symmetric membership functions are being used.

C. Framework for Adaptive Self-Diagnosis

The approach presented above provides a generic tool for self-diagnosis in core networks and specific service flows. The QoS diagnosis scheme though is static and defined in the initial configuration of the network management system. This is a major drawback due to the fact that, the environment conditions might change so the environment perception should change as well. Thus, the incorporation of a scheme for the evolution of the environmental perception is proposed as well.

The evolution of the diagnosis scheme is based on [4], where the incorporation of a supervised learning framework is being proposed. Such framework is based on a two layer approach where initially the ground truth is being built by using simple classification approaches (k-nearest neighbors - kNN), and then, by using this ground truth, the available measurements are used for the knowledge extraction. The main deficiency of such scheme is that the measurements are used for the extraction of the new membership functions in a simplistic manner using a well-known clustering method, the k-Means. In this paper we propose an approach which will take into account the diversity of the input measurements via statistical analysis of the monitored instances. Furthermore, we omit the training part used in [4], because the ground truth building procedure is time consuming, and requires a huge amount of training data. The considered approach here assumes an operational phase where data are being gathered on-the-fly and are then used for the adaptation of the input membership functions of the diagnosis mechanism.

In brief, the network management system monitors the network and proceeds in self-diagnosis. Each monitored tuple {Delay, Jitter, Packet Loss} is being evaluated by the fuzzy reasoner and classified as low, medium and high. In the low case the self-diagnosis mechanism identifies a problematic situation (in the case where the identifier mistakenly considers a situation as low QoS whereas it is not, the problem solving mechanism that undertakes to solve the identified problem intervenes without being needed -true negative). In the medium QoS case the self-diagnosis mechanism concludes that the information flow experiences medium QoS level, which is not problematic, but it could lead to either low or high QoS state without major alterations in the inputs. In the high QoS state the self-diagnosis mechanism identifies a normal situation. In the problematic case, where the self-diagnosis mechanism identifies a high QoS situation instead of
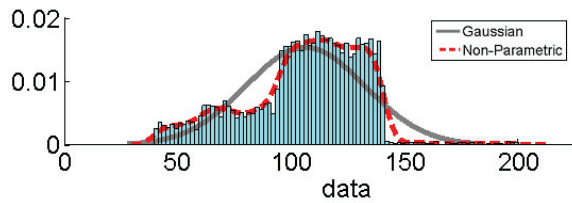
Figure 1: Gaussian and non-parametric statistical analysis for a given dataset

a low one, the problem solving mechanism does not intervene (false positive).

Once we have gathered enough (classified) measurements we have three sets of tuples, labeled as low, medium and high; misclassified tuples of the diagnosis mechanism from the true negatives and false positives are also included in the three sets. The classification is based solely on the current understanding of the decision maker on what constitutes low, medium and high respectively. The approaches that we have followed regarding the statistical processing of the measurements are two; the use of the Gaussian distribution and the non-parametric one (i.e., which uses the Kernel Density Estimator (Gaussian Kernel is used) of the measurements histogram) (Figure 1). The former approach is simpler whereas the latter should provide a better "fitting" to the available data.

For the Gaussian distribution approach, we obtain the mean value and the variance of each of the three states of each input (i.e. low, medium, high for Delay, Jitter and Packet Loss respectively). This enables the extraction of a Gaussian distribution as shown in Figure 1. The mapping of the Gaussian distribution to membership functions is straightforward so as to maximize the "correct" decisions and minimize the "wrong" ones.

For the non-parametric approach, we extract the density of the dataset in every point of the domain of definition and use the Kernel Density Estimator of the measurements' histogram for building the non-parametric curves. Then we normalize and map these curves into membership functions (Figure 1). As in the Gaussian case, the identification of the new membership functions is the adaptation mechanism so as to maximize the "correct" decisions and minimize the "wrong" ones.

III. PERFORMANCE ANALYSIS

In order to quantify the benefits from the introduction of the proposed scheme we have conducted a series of MATLAB simulations for the evaluation of the adaptive self-diagnosis scheme[1]. The algorithm of [4], which uses k-Means clustering for the identification of the interrelations among the inputs, is used as the benchmark/baseline algorithm. The dataset reflects small topologies and is built using generic {Delay, Jitter, Packet Loss} values available in the literature (Section II.B).

---

[1]The full dataset, accompanied by the source code is available at the http://kandalf.di.uoa.gr/IM2013/

We used 50.000 tuples; each tuple was described by three variables indicating the status of a measurement at time t_x, namely Delay ranging in [0…200] (ms), Jitter ranging in [0…200] (ms) and Packet Loss in [0…100] %. The dataset has been generated using typical values available in the literature for the VoIP service [8][9][10] using the normal distribution.

Initially, we configure the QoS event diagnosis mechanism with a generic membership function configuration and we compare with the ground truth. The ground truth is extracted using a fuzzy reasoner built specifically for the environment that is suitable for the generated inputs. Using a generic configuration the self-diagnosis scheme achieves a success rate of 64% (Table I). Given the fact that such self-diagnosis scheme is built to operate adequately in all environments we consider this success rate as acceptable. Then, we apply the benchmark adaptive algorithm [4] in this self-diagnosis scheme and achieve a success rate of 70.01% (amelioration of 8.6%) (Table II).

Table I: Initial configuration of the input membership functions of the self-diagnosis fuzzy reasoners

	Low	Medium	High
Delay (in ms)	0 – 80	40 – 150	>120
Jitter (in ms)	0 – 40	20 – 80	>60
Packet Loss (%)	0 – 0.0035	0.0025 – 0.008	>0.006

Table II: Input membership functions of the self-diagnosis fuzzy reasoners after the clustering adaptation procedure

	Low	Medium	High
Delay (in ms)	0 –20	5 –80	30 –200
Jitter (in ms)	0 – 40	0.35 – 1	0.55 –2
Packet Loss (%)	0 – 0.005	0.004 – 0.0057	0.0055 – 0.01

By incorporating the adaptive mechanism with the Gaussian approach and following the methodology presented in Section II.C we modify the input membership functions as captured by Table III. The input states are captured by new membership functions, which are described by Gaussian distributions, with higher overlap areas. The success rate of the adapted scheme reaches 84.072% compared to the ground truth (an amelioration of 23.7%). The required time for the processing of the dataset and the extraction of the new membership functions is 13.07 seconds in an average consumer laptop (i.e. Quad core, 1.6 GHz, 4 GB RAM).

For the same dataset, we also apply the non-parametric approach. Given the fact that for the adaptation of the membership functions we must have a finite number of points (MATLAB fuzzy logic toolbox limitation [11]) we choose 16 points of the extracted distribution curves and we define the membership functions. The new membership functions are closer to the actual distribution of the dataset (Figure 2) and reach a success rate of 84.16% compared to the ground truth (an amelioration of 23.95%). In this case the required time for
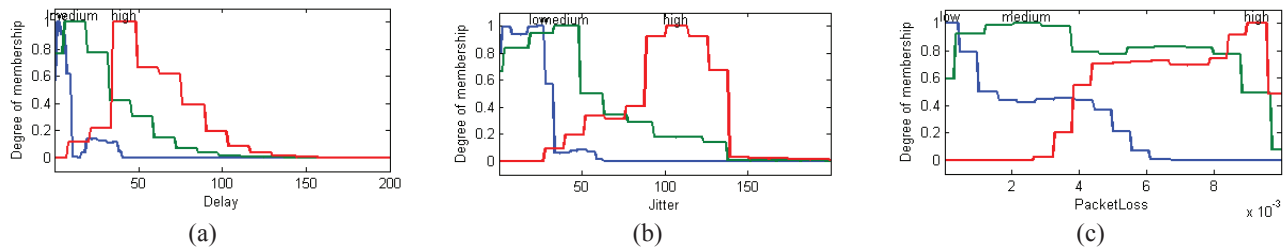
Figure 2: Input (Delay – (a), Jitter – (b), and Packet Loss (c)) membership functions after the non-parametric adaptation procedure.

the adaptation procedure (processing and extraction of the new membership functions) is 22.38 seconds.

Thus we observe that the introduction of the newly developed and more sophisticated schemes is beneficial for the network, given the fact that the enhanced mechanisms have better success rates compared to the initial (benchmark algorithm). Furthermore, we observe that both algorithms, Gaussian and non-parametric one, have similar results (~84% success rate) with the latter requiring approximately 40% more processing time compared to the Gaussian one (the required time is indicative of the processing cost of the proposed solution and should not be considered as an absolute value taking into account that the validation is performed using MATLAB).

Table III: Mean values (μ) and standard variations (σ) of the input membership functions of the Gaussian adaptation scheme

		Low	Medium	High
Delay (in ms)	μ	9.8	30.32	64.67
	σ	10.48	22.2	22.81
Jitter (in ms)	μ	18	48	106
	σ	12	34	26.11
Packet Loss (%)	μ	0.0022	0.0047	0.0073
	σ	0.0017	0.0028	0.0018

IV. CONCLUSION

This paper proposes a scheme for adaptive self-diagnosis in core network elements. The solution is based on a fuzzy logic reasoner which adapts the environment interpretation based on the statistical analysis of the monitored inputs, using two approaches; the Gaussian and the non-parametric one. The fuzzy reasoner aims at identifying events by using network measurements and service information in order to proceed in advanced decisions. Both approaches proposed in this paper have been validated through MATLAB simulations by applying them to a specific case study for core network elements and more specifically for VoIP QoS degradation events' identification. The performance analysis revealed that the adaptation framework significantly improves the performance of the event identifier. Both approaches (i.e., Gaussian and non-parametric) outperform significantly the benchmark algorithm and lead to much higher event identification rate (~84% instead of ~70%, an amelioration of 20%). Comparing the two proposed approaches, the Gaussian and the non-parametric one, we should highlight the fact that even though the latter is closer to the distribution of the dataset and theoretically is more suitable to the available data, both

achieve almost the same identification rates. Furthermore, the non-parametric one has significantly higher processing cost which in our MATLAB simulation is captured by almost 40% more time required for the non-parametric adaptation compared to the Gaussian approach. Such additional processing overhead could become an issue of importance in larger topologies and/or when a larger number of inputs are used for the identification process.

Our future work includes the incorporation of Support Vector Machines (SVMs) for the classification of the data; however, such approach might increase the complexity of the algorithm therefore the trade-off between complexity and identification success rate will need to be evaluated.

REFERENCES

[1] B. Yan, G. Chen, "Model-based fault Diagnosis for IEEE 802.11 wireless LANs", MobiQuitous '09. 6th Annual International Mobile and Ubiquitous Systems: Networking & Services, 2009.

[2] J. de Kleer, B. C. Williams. "Diagnosis with behavioral modes", In Proceedings International Joint Conferences on Artificial Intelligence, pages 1324–1330, Detroit, MI, Aug. 1989.

[3] Yong Sheng, Tan, K., Guanling Chen, Kotz, D., Campbell, A., "Detecting 802.11 MAC Layer Spoofing Using Received Signal Strength," IEEE 27th Conference on Computer CommunicationsINFOCOM 2008. pp.1768-1776, 13-18 April 2008

[4] Panagis Magdalinos, Apostolos Kousaridas, Panagiotis Spapis, Giorgos Katsikas, Nancy Alonistioti, "Feedback-based Learning for Self-Managed Network Elements Decision Making Evolution", 12th IFIP/IEEE International Symposium on Integrated Network Management, Dublin Ireland, May 2011.

[5] Oliver Johnson, "Information Theory And The Central Limit Theorem" , Imperial College, 2004, ISBN 1-86094-473-6

[6] Na Luo, Feng Qian, "Evolutionary algorithm using kernel density estimation model in continuous domain," Asian Control Conference, 2009. ASCC 2009. 7th, pp.1526-1531, 27-29 Aug. 2009

[7] P.Magdalinos et al "Coverage and Capacity Optimization of Selfmanaged Future Internet Wireless Networks", in ServiceWave, pp 201-202, 2010

[8] Quality of Service for Voice over IP, CISCO, 6/30/2001

[9] Latency and QoS for Voice over IP, SANS Institute 2004

[10] http://www.slac.stanford.edu/comp/net/wan-mon/tutorial.html

[11] MathWorks, MATLAB, Fuzzy Logic Toolbox™ User's Guide, R2012a