A Cognitive Quality of Transmission Estimator for Core Optical Networks

Tamara Jiménez, Juan Carlos Aguado, Ignacio de Miguel, Ramón J. Durán, Marianna Angelou, *Member, IEEE*, Noemí Merayo, Patricia Fernández, Rubén M. Lorenzo, Ioannis Tomkos, and Evaristo J. Abril

***Abstract*—We propose a cognitive Quality of Transmission (QoT) estimator for classifying lightpaths into high or low quality categories in impairment-aware wavelength-routed optical networks. The technique is based on Case-Based Reasoning (CBR), an artificial intelligence technique which solves new problems by exploiting previous experiences, which are stored on a knowledge base. We also show that by including learning and forgetting techniques, the underlying knowledge base can be optimized, thus leading to a significant reduction on the computing time for on-line operation. The performance of the cognitive estimator is evaluated in a long haul and in an ultra-long haul network, and we demonstrate that it achieves more than 98% successful classifications, and that it is up to four orders of magnitude faster when compared with a non-cognitive QoT estimator, the Q-Tool.**

***Index Terms*—Case-based reasoning (CBR), cognitive networks, impairment-aware networking, quality of transmission (QoT), wavelength-routed optical network (WRON).**

I. Introduction

In all-optical transparent networks, traffic is carried through end-to-end wavelength channels, called lightpaths, without any type of Optical-Electrical-Optical conversion at intermediate nodes. However, as optical signals traverse fiber links and nodes, and propagate through active and passive optical components towards their destination, they suffer from a number of physical impairments which degrade the signal quality. These impairments affect each optical channel individually, but they also cause disturbance and interference between co-propagating channels. Hence, as there is no conversion to the electrical domain and consequently, no regeneration at intermediate nodes, the Quality of Transmission (QoT) will be affected and might not comply with service requirements. Therefore, in the last years, the impact of physical layer impairments in optical network design and operation has received significant attention. As a result, this interest has led to a set of proposals that, for instance, not only solve the Routing and Wavelength Assignment (RWA) problem in Wavelength-Routed Optical Networks (WRONs), but also ensure appropriate QoT on the established lightpaths [1], [2]. For that aim, effective and efficient methods for predicting the QoT of lightpaths (before being established and measured) are required. In that way, such a predicting tool can be used to discard those lightpaths that will not fulfill QoT requirements, and also to verify that a new lightpath will not have a significant impact on existing ones, thus avoiding troublesome situations.

In particular, Azodolmolky *et al.* [3], [4] have presented an impairment aware network planning and operation tool for all-optical and translucent networks. A key element of that tool is an integrated real-time quality of transmission estimator, the Q-Tool. This tool combines in a single framework a number of well investigated and verified analytical models previously proposed in the literature and, in contrast to other approaches, it also relies on a numerical split-step Fourier method in order to improve accuracy [3]. The Q-Tool receives the topology (with its physical characteristics) and a set of lightpaths and then computes their associated Q-factors. The Q-factor is an indicator of the quality of transmission, which is related to the signal's bit error rate (BER), so that a higher value of the Q-factor corresponds to a lower BER [5]. The Q-Tool provides relatively accurate estimates of the Q-factor by taking into account several models of linear and nonlinear impairments of the physical layer, thus being a very useful element in optical network design and control. However, it suffers from a few limitations. First of all, it is only valid for 10 Gb/s OOK networks. Secondly, due to the complex calculations required, the computing time is very high, ranging from 1 to 1,000 seconds, depending of the scenario, on the software implementation described in [6]. Therefore, the use of this tool may be prohibitive when time constraints are important for real time control; as well as for the application of a number of planning techniques, such as those based on genetic algorithms [7], since they rely on the evaluation of many alternative potential configurations.

On the other hand, Poggiolini [8] has proposed a Gaussian Noise (GN) model which is able to estimate, quickly and accurately, the Optical Signal to Noise Ratio (OSNR) of the optical channels in uncompensated coherent transmission systems. Although this pioneering work opens the door for further developments and enhancements, it does not yet address network scenarios (where channels coming from different locations are multiplexed in an optical fiber at optical cross-connects), and it is not valid for dispersion compensated systems.

In this paper, we propose an alternative approach for predicting the quality of transmission of lightpaths in an optical

Manuscript received October 22, 2012; revised December 11, 2012; accepted December 28, 2012. Date of current version January 30, 2013. This research has been supported by the CHRON (Cognitive Heterogeneous Reconfigurable Optical Network) project, with funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 258644. The work of T. Jiménez was supported in part by the Council of Education of the Regional Government of Castilla-León and in part by the European Social Fund.

T. Jiménez, J. C. Aguado, I. de Miguel, R. J. Durán, N. Merayo, P. Fernández, R. M. Lorenzo, E. J. Abril are with the University of Valladolid, 47011 Valladolid, Spain (e-mail: tamara.jimenez@tel.uva.es).

M. Angelou and I. Tomkos are with the Athens Information Technology Center, 19002 Peania, Athens, Greece.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2013.2240257

network (i.e., before being established), which consist on relying on cognition. Thus, by exploiting previous experiences (which are stored on a knowledge base), fast and correct decisions on whether a lightpath fulfills QoT requirements or not, can be made without having to rely on complex methods. In particular, we propose a novel cognitive technique, based on Case-Based Reasoning (CBR) [9], which provides successful classifications of lightpaths into high or low QoT categories in more than 98% cases, and that is several orders of magnitude faster than when using the Q-Tool; thus becoming a promising technique for highly dynamic impairment-aware optical networks.

The focus of this paper is set on the description of this novel technique, and on demonstrating the advantages of the use of cognition for QoT estimation. Since solid work on QoT estimation already exists for 10 Gb/s OOK networks, we have targeted the analysis of these scenarios in order to have a reliable baseline method for our comparisons. Considering this, we have selected the Q-Tool, as it gets a good balance between accuracy (considering most of the physical impairments) and speed, thereby being a very good option to showcase the advantages and potential of the cognitive estimator. On the other hand, the fundamentals of the cognitive technique are generic enough to be applied to networks with higher data rates and, for instance, we have recently demonstrated its application in a WDM 80 Gb/s PDM-QPSK system test bed [10]. Therefore, it should be noted that the application of cognition to quality of transmission assessment is the novel contribution of this paper, being the Q-Tool just used as a benchmark method for evaluating the capabilities of the cognitive mechanism under conditions as realistic as possible.

The remaining of this paper is organized as follows. First of all, in Section II, we explain the fundamentals of the cognitive QoT estimator, which was introduced in [11], and discuss how the underlying knowledge base can be optimized by means of learning and forgetting techniques [12]. Then in Section III, the performance of the cognitive QoT estimator is analyzed by means of a simulation study on a long-haul and in an ultra-long haul network, with different numbers of nodes, in order to analyze potential scalability issues. Moreover, we also discuss how the initial knowledge base can be built in pragmatic networking scenarios. Finally, in Section IV, the main conclusions are stated.

II. Description of the Cognitive QoT Estimator

We have developed a cognitive QoT estimator which classifies lightpaths into two categories: high and low QoT. These categories are determined by means of a user-defined threshold on the Q-factor ($Q_{threshold}$). Thus, if the Q-factor of a lightpath is higher or equal to this threshold, then it belongs to the high QoT category, and we assume that the lightpath complies with quality requirements. Otherwise, the lightpath belongs to the low QoT class. Alternatively, the classification can be done according to other QoT parameters like Error Vector Magnitude (EVM) values [5], [10], in this case being the values lower than the threshold those associated to the high QoT class. Nevertheless, in this paper we focus on the Q-factor as the parameter determining the category of each lightpath.

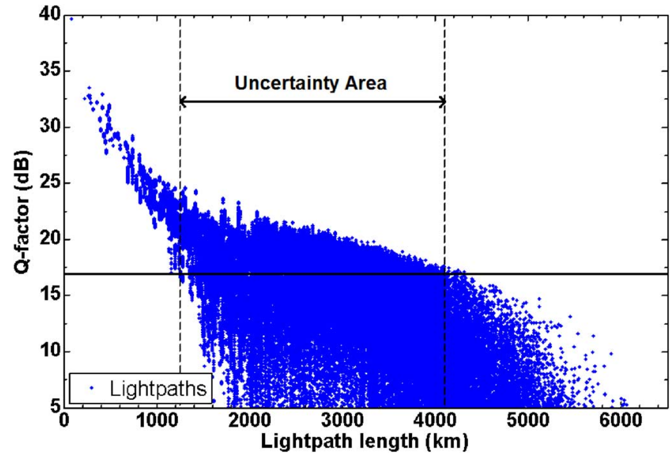

Fig. 1. Q-factor of the lightpaths, as a function of their lengths, in a simulation of the GÉANT2 network for different network loads and 32 wavelengths per link.

The cognitive estimator determines the QoT category of a lightpath by means of a hybrid mechanism. First, it takes into account the length of the lightpath, and then employs, if necessary, a Case-Based Reasoning system. The motivation for using the length as a first element to classify the lightpath is its significant impact on the Q-factor.

As an example, Fig. 1 shows the Q-factor of different lightpaths depending on their total length. The figure has been obtained when simulating the GÉANT2 network [13] as a dynamic WRON, equipped with 32 wavelengths per link and 10 Gb/s OOK transceivers, for different network loads, and using the MATLAB implementation of the Q-Tool to evaluate the Q-factor of the lightpaths. The threshold to distinguish between high and low QoT categories has been set to 16.9 dB (which corresponds to a BER of 10^{-12}). As it can be observed, lightpaths shorter than a certain length ($\sim$1,250 km) generally belong to the high QoT category, while those with very long lengths ($\sim$4,100 km) typically belong to the low QoT class. However, there is an uncertainty area (i.e., a range of intermediate lengths) where the rest of the characteristics of the lightpath also play an important role on its Q-factor and hence on its classification into a QoT category. Therefore, for classifying the lightpaths in this uncertainty area, a CBR system is applied.

A. Case-Based Reasoning for QoT Assessment

CBR is a problem solving paradigm that emphasizes the role of prior experiences or cases stored in a Knowledge Base (KB) [9]. In CBR, a new problem is solved:

- by retrieving from the KB the most similar cases faced in the past to the problem currently tackled,
- by reusing the retrieved cases, either directly or adapting them in order to provide a solution for the new problem,
- by revising the proposed solution, i.e., by checking its performance,
- and by (maybe) retaining in the KB the new case and the solution employed.

These steps can be implemented in different ways depending on the final application and the limiting factors like the maximum time to provide the solution or the desired precision of the solution. An excellent review of the main techniques existing to implement a CBR system can be found in [14].

Particularly, in the CBR system of the cognitive QoT estimator, the initial KB is composed by a number of cases, which consist of a description of the lightpath (i.e., a set of attributes) and its associated Q-factor. The description of the lightpath contains its route, that is, the set of links that it traverses (represented by the percentage of their individual contribution to the total length of the lightpath), the selected wavelength, its total length, the sum of the co-propagating lightpaths per link, and the standard deviation of the number of total co-propagating lightpaths. Moreover, the associated Q-factor stored in the KB is an estimate of the quality of transmission which has been obtained by using the Q-Tool. In order to obtain these cases, previous off-line simulations are executed. Therefore, the cases in the KB are different lightpaths established at different moments of those simulations, and their associated Q-factors are calculated off-line by means of the Q-Tool.

The reason for selecting the Q-Tool along this study is that, as explained in the introduction, it combines several verified physical layer impairment models and numerical computations in order to provide more accurate values than other proposals. Nevertheless, it is important to note, that although we have used the Q-Tool for populating the KB, any other QoT estimator could be used instead and, in fact, it could be populated with data obtained from an optical communications system simulator or even coming from optical network monitors [15]. (In Section III-B we will provide additional insights about these alternative procedures for building the KB.)

In real network operation, where fast assessment of lightpath quality is required, the cognitive QoT estimator works as follows. First, when facing a new lightpath request, the RWA problem is solved and the total length of the lightpath is calculated. If the length is lower than the lower threshold of the uncertainty area, the lightpath is assumed to fulfill the QoT requirements. On the other hand, if the length is higher than the upper threshold of the uncertainty area, it is assumed that the QoT requirements are not fulfilled. However, if the length belongs to the uncertainty area, the CBR system is applied and it retrieves the most similar lightpath from the KB to the new request.

In order to assess the similarity when comparing the new lightpath (x) with each one contained in the KB (y), the attributes are normalized and the weighted Euclidean distance is calculated [16], [17] following (1),

$$Similarity\,(x, y) = -\sqrt{\sum_{a=1}^{n} W_a^2 \cdot (x_a - y_a)^2} \qquad (1)$$

where a represents each attribute of the lightpaths x and y, W_a is the weight associated to that attribute, and n is the set of attributes. Thus, higher values (i.e., closer to zero values) of (1) mean higher similarity of the cases. The set of weights used are the least-squares regression coefficients of a linear model for the KB considering the Q-factor as the dependent variable.

The Q-factor of the new lightpath is assumed to be the same one than that of the retrieved case (i.e., the most similar lightpath in the KB), and that value is used to decide whether the lightpath fulfills the QoT requirements or not by comparing with the threshold value $(Q_{threshold})$.

In this first version of the cognitive QoT estimator, the KB is completely static, and so it is not updated with new cases nor optimized, (i.e., the retain stage of CBR is not used). This first version of the cognitive QoT estimator will be denoted as R-CBR (Regular-CBR).

B. *Optimization of the KB*

As it was previously mentioned, the KB of a CBR system can be updated to include new experiences by storing the description and solutions of new problems faced by the CBR system, i.e., the CBR system can learn. Learning tends to increase the effectiveness of the system, as the KB grows. However, excessive learning has a great impact on retrieval time, which is incremented, as it strongly depends on the size of the KB [14]. This is known as the utility problem [14], [18]: the cost of maintaining and searching in a large case base outweighs the benefit of storing its knowledge.

Therefore, to avoid the utility problem, not only learning but also forgetting techniques have to be implemented. Thus, case addition and deletion strategies should be implemented to control retention and to eliminate cases that do not improve the performance of the system. Hence, in this subsection, we propose a mechanism to enhance the cognitive QoT estimator with the execution of periodic maintenance stages where the KB is updated and optimized.

During the operation of the cognitive estimator, all cases whose classifications have been done by the CBR stage (i.e., lightpaths with lengths belonging to the uncertainty area) follow a double check. It is checked (1) if the new lightpath has been correctly classified in its QoT category, and (2), if the error between the Q-factor estimate obtained by the CBR system and its real value is below a certain amount (the permitted error, $\varepsilon_{permitted}$). If (1) or (2) are not fulfilled, then the case is stored in an auxiliary database as a candidate to be incorporated to the KB (i.e., to be learned).

When the CBR system has made a certain number of classifications (either correct or wrong), the maintenance phase is executed. It consists in, first, adding to the KB the cases stored as candidates to be learned (and then resetting the auxiliary database), and second, applying a technique to remove redundant cases from the KB.

Specifically, the technique selected to carry out the redundancy removal is based on the Conservative Redundancy Reduction (CRR) method [19]. This algorithm aims at removing redundant cases which are not located near the class borders. To do this, the coverage set (CS) of each case is calculated. The coverage set of a case (c) is the set of all cases that c can successfully classify [19], [20]. Therefore, cases which have a large CS are probably situated in clusters of cases with the same classification. On the other hand, if a case has a small CS, this indicates that it has few neighbors, and therefore, it is situated close to a border of the class [19]. The pseudo code to calculate the CS of each of the cases in the KB, adapted to the cognitive QoT estimator features, is shown in Table I.

TABLE I
PSEUDOCODE TO CALCULATE THE COVERAGE SET OF ALL CASES IN THE KB

```
For each case c in the KB {
  Set CS(c)= ∅;
}
For each case q in the KB {
  1: c ← next nearest neighbor of case q;
  if (c predicts q properly) {
    CS(c) ← CS(c) ∪ q;
    Go to 1;
  }
  else next q;
}
```

Note: c predicts q properly if c classifies q into the correct category (i.e., if they both belong to the same class), and if the absolute error committed in the prediction of the Q-factor for q (i.e., the difference between the real Q-factors of q and c, in absolute value) is lower than $\varepsilon_{permitted}$.

Once the CS has been calculated, the CRR algorithm sorts all cases in the KB in ascending order according to the size of their coverage set. Then, the cases in the KB are analyzed starting from that with the smallest CS, and the cases in its coverage set are removed from the KB [19]. (Obviously, if a case is removed during this process, it will not be analyzed later to delete its own coverage set.)

If after running these processes the size of the KB is higher than that of the original KB (i.e., the KB that was generated before starting the whole optimization process), then an appropriate number of cases is deleted from the KB, starting with those cases having a higher coverage set. Therefore, the resultant size of the KB is never higher than its initial size.

Finally, it should be noticed that in order to apply the optimization procedure that we have just described, it is necessary to compare the estimated Q-factor for each lightpath with the real one. Hence, for real-time updates of the KB, the cognitive QoT estimator must work in collaboration with a network monitoring system (which measures the Q-factors of the established lightpaths) and supported by appropriate network protocols. Nevertheless, in this paper, we focus on the off-line optimization of the KB. Hence, once an initial KB is generated by means of off-line simulations (as described in Section II-A), it is optimized by applying the procedure mentioned above. This is also done by means of an off-line simulation (i.e., executed prior to online operation) which uses the Q-Tool to provide the "real" Q-factors.

Summing up, we have proposed two different methods. The first one, referred as R-CBR (Regular CBR), is a cognitive QoT estimator that does not optimize the KB prior to online operation (i.e., it operates as described in Section II-A) [11]. The second one, called FixE-CBR (Fixed Error CBR), is a cognitive estimator which applies learning and forgetting techniques in order to perform an off-line optimization of the KB with a fixed permitted error ($\varepsilon_{permitted}$) [12]. However, the KB associated to the FixE-CBR method is no further optimized during online operation.

III. SIMULATION SCENARIOS AND RESULTS

A. Performance Evaluation of the Cognitive QoT Estimator

To evaluate the performance of the two versions of the cognitive QoT estimator, simulations have been carried out in two different networks in order to analyze potential scalability issues: a long haul network, the 14-node Deutsche Telekom (DT) network [3], and an ultra-long haul network, the 34-node GÉANT2 network [13]. Both networks have been configured as dynamic WRONs and equipped with 10 Gb/s OOK transceivers. Each link consists of a number of spans formed by Standard Single Mode Fiber (SMF) and Dispersion Compensating Fiber (DCF), and 32 and 64 wavelengths per link have been considered. The results have been obtained by analyzing the networking scenarios under different traffic loads, and the traffic loads for the DT and GÉANT2 networks have been selected so that they lead to a similar range of blocking probabilities. For the DT network, traffic loads from 0.3 to 2.0 for the 32 wavelengths scenario, and from 0.5 to 4.1 for the 64 wavelengths scenario, have been considered. For the GÉANT2 network, the considered network loads have been 0.1 to 0.45, and 0.1 to 1.0, for the 32 and 64 wavelengths scenarios, respectively. A traffic load of 1 means that, in average, if there were no blocking, there would be one lightpath established between each source-destination pair in the network. The routes and wavelengths for the connections have been obtained by means of an adaptive RWA algorithm (AUR-Exhaustive) [21], since it offers more flexibility and thus a much lower blocking probability in dynamic scenarios than other approaches based on the utilization of fixed pre-calculated routes [21], [22].

TABLE II
LOW AND HIGH LENGTH LIMIT OF THE UNCERTAINTY AREA

Network	Number of wavelengths	Low length limit (km)	High length limit (km)
DT	32	975	1,875
	64	975	2,050
GÉANT2	32	1,250	4,125
	64	1,175	4,225

Two implementations of the Q-Tool were developed in the framework of the European Union DICONET project [3], [4]: a MATLAB implementation, and a hardware implementation (based on an FPGA) which accelerates the QoT estimation process [6]. With the aim of providing a fair comparison in terms of computing time, the cognitive QoT estimator has also been implemented in MATLAB, and thus it will be compared with the software implementation of the Q-Tool.

The Q-factor threshold for the classification of lightpaths into high or low QoT categories ($Q_{threshold}$) has been set to 16.9 dB (i.e., distinguishing between BER lower and higher than 10^{-12}, respectively). Moreover, by means of simulation, the limits (in terms of lightpath length) of the uncertainty area have been determined so that the probability of successful classification is higher than 99.99% outside the uncertainty area. The limit lengths for both networks can be found in Table II.

The initial KB of the CBR system has been populated with different numbers of cases, running from 500 to 5,000 for the DT network and from 5,000 to 50,000 for GÉANT2. In order to ensure a fair comparison between both networks, the size of the KBs involved in the GÉANT2 network has been increased, since it has a higher number of nodes involved (the number of source-destination pairs increases by 6 times from the DT to the GÉANT2 network). The cases of the KB for both networks have been chosen randomly from those generated in an off-line simulation. Each KB provides coverage of the uncertainty area for

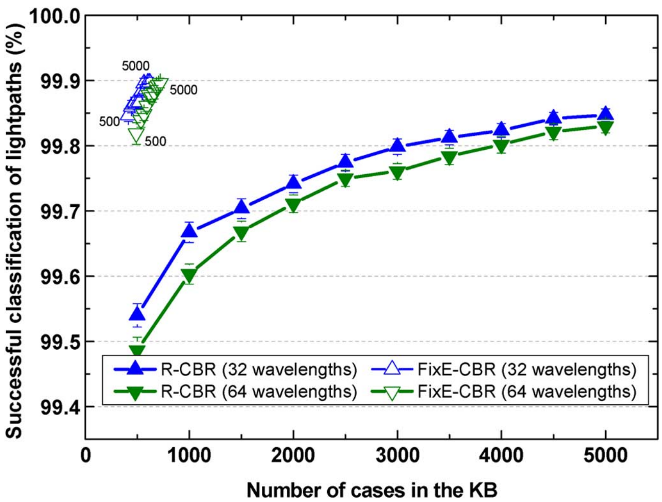

Fig. 2. Successful classifications of QoT when comparing R-CBR and FixE-CBR methods for DT network. The numbers that appear next to the FixE-CBR points refer to the initial sizes of the KB before executing the optimization procedure.

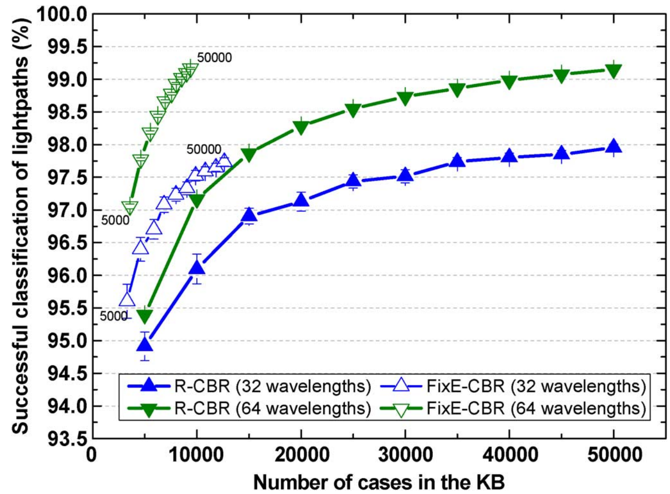

Fig. 3. Successful classifications of QoT when comparing R-CBR and FixE-CBR methods for GÉANT2 network. The numbers that appear next to the FixE-CBR points refer to the initial sizes of the KB before executing the optimization procedure.

all traffic loads, so that the same KB can be used independently of the traffic load faced by the network.

When operating with an optimized KB (i.e., for FixE-CBR), an off-line KB optimization process has been executed. For that objective, 6,000 new lightpaths for DT and 36,000 for GÉANT2, belonging to the uncertainty area, have been classified, running the optimization process described in Section II-B after every 500 classifications. The permitted error ($\varepsilon_{permitted}$) when optimizing the KB has been set to 3 dB.

Once the optimization process has finished, the performance of the cognitive QoT estimator is analyzed. For that aim, other 6,000 lightpaths in the DT network and 36,000 in GÉANT2 network (belonging to both the certainty and uncertainty areas) have been evaluated. However, the KB is no longer updated during this evaluation, i.e., there is no additional learning during the evaluation.

The results that we show in the following figures have been obtained after repeating this process 100 times with different KBs. Average results are represented together with 95% confidence intervals (although in most cases the size of the confidence intervals is smaller than the size of the symbols).

Fig. 2 represents the percentage of successful classifications of lightpaths into high or low QoT categories for the DT network when employing the R-CBR and FixE-CBR estimators, that is, it compares the successful classifications when the KB has not been optimized before online operation, and when it has been optimized with a fixed error policy. In the figure, the numbers written next to the points associated to the FixE-CBR method indicate the size of the initial KB (i.e., before being optimized).

As shown in the figure, even when the KB is not optimized (R-CBR), the percentage of successful classifications is very high. For the smallest size of the KB (500 cases), the cognitive QoT estimator achieves more than 99.45% correct classifications, and that percentage raises to 99.8% for the highest size of the KB considered in the simulation (5,000 cases).

When R-CBR and FixE-CBR are compared, not only is the percentage of successful classifications with an optimized KB (FixE-CBR) higher than without optimization (R-CBR), but also the number of cases in the final KB is typically much lower with FixE-CBR. For example, for 32 wavelengths, and a KB with an initial size of 500 cases and a final size of 412 cases, the percentage of successful classifications raises from 99.53% to 99.84%. On the other hand, for an initial size of the KB equal to 5000 cases, and again for the 32 wavelengths scenario, FixE-CBR slightly raises the percentage of successful classifications from 99.84% (R-CBR) to 99.89% and, more importantly, it achieves a significant reduction of the size of the KB, as the final KB only contains 618 cases versus the 5,000 initial cases (i.e., 87.64% reduction). As we will demonstrate later, this fact has a significant impact in terms of reducing the computing time.

For the GÉANT2 network, a similar behavior can be observed. Fig. 3 compares the evolution of the percentage of successful classifications when the KB size is increased for R-CBR and FixE-CBR for this network. As it can be seen, for the 32 wavelengths scenario, the highest percentage of successful classifications reaches 98% for a KB of 50,000 cases, and 99.15% when considering 64 wavelengths and the same size of the KB. Moreover, FixE-CBR improves the percentage of successful classifications for small KB sizes. For example, for 64 wavelengths and an initial KB size of 5,000 cases, the percentage raises from 95.4% to 97%. However, again, the greatest impact of FixE-CBR is the significant reduction in the size of the KB. In this way, for 64 wavelengths, Fix-CBR reduces the KB size from 50000 to 9404 cases (i.e., 81.19% reduction).

These results seem to indicate that there is a scalability problem, as the results for the GÉANT2 network are slightly worse than for the DT. Therefore, we have analyzed this issue in more detail.

The cognitive estimator relies on a hybrid mechanism that first decides by means of a threshold length, and then, if required, by a CBR system. As previously described, the threshold lengths have been set so that when a classification

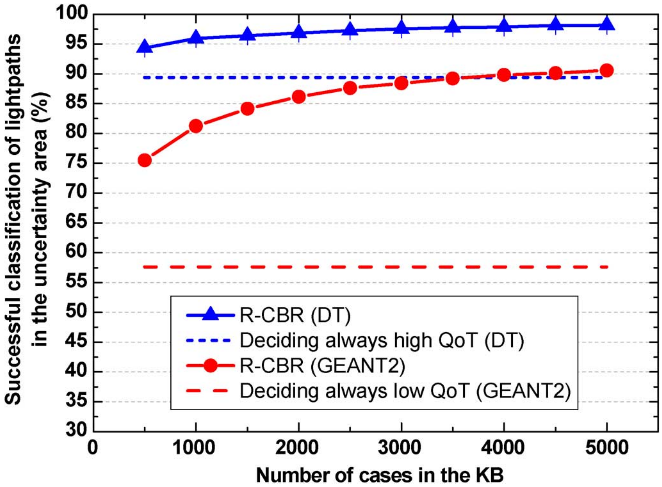

Fig. 4. Successful classifications of CBR in the uncertainty area for DT and GÉANT2 networks with 32 wavelengths and successful classifications in the uncertainty area betting for the most likelihood category.

can be done according to them, the ratio of successful classifications is higher than 99.99%. For the DT network most of the cases are solved by means of the threshold mechanism and only the 8% cases are solved by the CBR system (for both 32 and 64 wavelengths). In contrast, for the GÉANT network, the percentage of lightpaths that have to be solved by the CBR stage raises to 53.8% and 56% for 32 and 64 wavelengths, respectively. Therefore, the GÉANT network scenario poses a more difficult problem to solve as the length of the lightpaths is not as a determinant factor as in the DT network.

To further analyze this issue, Fig. 4 shows the percentage of successful classifications when the CBR stage is used (i.e., when lightpaths to be classified belong to the uncertainty area). It is worth noting that the size of the KB has been set to the same values for both networks, in order to facilitate a direct comparison. As depicted in Fig. 4, for a KB of 5,000 cases, the percentage of successful classifications reaches up to 97.5% for the DT network and 90.56% for GÉANT2. This difference between the percentages is the result of the dependence of the distribution of cases into the two categories. Thus, if we implement a QoT estimator which decides that all the lightpaths belong to most probable category for the DT network (i.e., a majority class classification), the percentage of successful classifications in the uncertainty area would reach up to 89.5%, whereas for the GÉANT2 network the percentage would only reach up to 57.59%. Therefore, the CBR mechanism improves the results by only 8 percentage points for the DT network when compared with the majority class classification, while for the GÉANT2 network it does by 32.97 percentage points. As a conclusion, the worst results of the cognitive QoT estimator for the GÉANT2 network when compared with the results for the DT network are not only due to the size of the network but also to the more complex structure of the data.

Next, we analyze the motivation and impact of implementing the QoT estimator as a hybrid system. The main reason for considering the length of the lightpath as a first parameter to make a decision is the reduction of computing time when compared with a CBR-only approach. For instance, for a KB of 5,000 cases, the mean time to classify a lightpath is 6.6 ms for the DT network, whereas the mean time for the hybrid approach with the same KB size is 0.5 ms. Thus, the first phase, where a decision may be made by considering the length of the lightpath, accelerates the process without degrading the performance in terms of successful classifications. However, relying exclusively on the length to classify the lightpaths into QoT categories (i.e., without using the CBR system), while extremely fast, is not a good method. To demonstrate this, Table III collects the percentage of successful classifications for the worst source-destination pair (i.e., that pair having the lowest percentage of successful classifications). The results have been obtained with the cognitive QoT estimator (R-CBR), and also with an estimator that makes decisions by only taking into account the length of the lightpaths. For the former, a KB of 5,000 cases has been considered for the DT network, and of 50,000 cases for the GÉANT2. For the latter, the threshold to classify lightpaths into high or low QoT categories was set, by means of simulations, to the threshold leading to the best results in terms of global successful classifications ratio. As it can be noticed, for the DT network, the threshold length-based estimator achieves a poor successful classification rate of 64% for 32 wavelengths, and 59.7% for 64 wavelengths, for the worst source-destination pair. In contrast, the cognitive estimator outperforms it, as its success ratio for the worst source-destination pair for both wavelengths is 96.1%. The same behavior is observed for the GÉANT2 network. It is noticeable that for this network, the threshold length-based estimator achieves even worse results. In fact, there are pairs of source-destination nodes whose associated lightpaths are always incorrectly classified. In contrast, the R-CBR method obtains successful classifications of 77% and 87% for 32 and 64 wavelengths, respectively.

TABLE III
SUCCESSFUL CLASSIFICATIONS OF LIGHTPATHS FOR THE WORST SOURCE-DESTINATION PAIR

Network	Number of wavelengths	R-CBR Successful classifications (worst case)	Threshold length based estimator	
			Threshold length	Successful classifications (worst case)
DT	32	96.13%	1,375 km	64.23%
	64	96.15%	1,425 km	59.77%
GÉANT2	32	77.05%	1,725 km	0%
	64	87.15%	1,725 km	0%

As we have mentioned, the use of learning and forgetting techniques leads to a reduction of the size of the KB, which translates in a lower computing time (i.e., the time employed to estimate the QoT of one lightpath). This can be noticed in Figs. 5 and 6, where the computing time (per lightpath) of the Q-Tool and the two cognitive QoT estimators are represented versus the sizes of the KBs for the DT and GÉANT2 networks, respectively, both equipped with 32 wavelengths. The simulation of both tools was run on a Debian GNU/Linux 6.0 machine using one core of an AMD Opteron 6128 processor.

It should be noted that, since in the dynamic operation of a network not only the QoT of each new lightpath to be established must be assessed, but also that of co-propagating ones

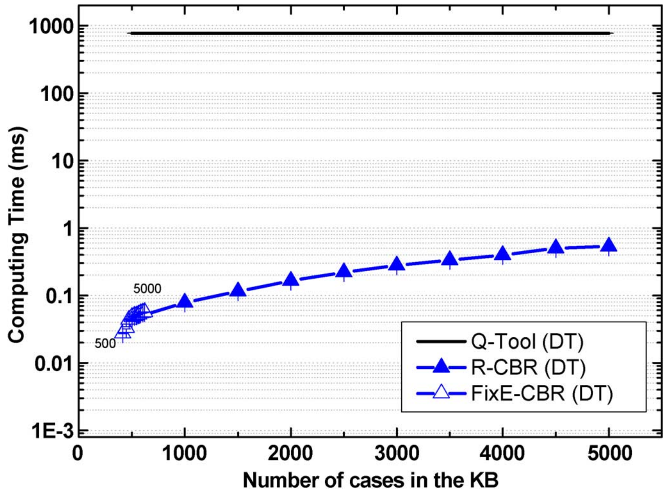

Fig. 5. Computing time to assess the QoT per lightpath for the DT network for 32 wavelengths using the Q-Tool, R-CBR and FixE-CBR methods. The numbers that appear next to the FixE-CBR points refer to the initial sizes of the KB before executing the optimization procedure.

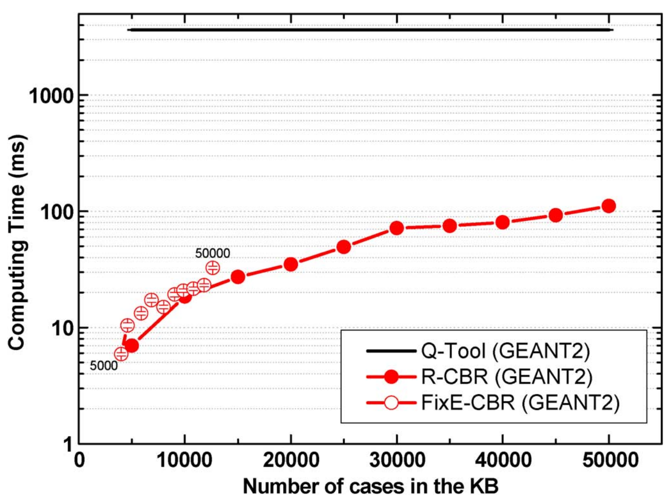

Fig. 6. Computing time to assess the QoT per lightpath for the GÉANT2 network for 32 wavelengths using the Q-Tool, R-CBR and FixE-CBR methods. The numbers that appear next to the FixE-CBR points refer to the initial sizes of the KB before executing the optimization procedure.

(to verify that they will not be disrupted by the new connection), a low computing time per lightpath is required, especially in highly dynamic networks.

As it can be observed, for the DT network, the basic cognitive QoT estimator (R-CBR) is around three orders of magnitude faster than the Q-Tool when the KB contains 5,000 cases. However, the computing time when the KB is optimized, is even lower, as the size of the final KB is significantly reduced. In particular, the computing time of FixE-CBR is close to four orders of magnitude lower than that of the Q-Tool.

For the GÉANT2 network, the computing time is higher, both for the Q-Tool (requiring approximately 3.6 s) and for the cognitive QoT estimator. Regarding the cognitive estimator, this increase is due to the fact that the GÉANT2 is a more complex network, with a higher number of links. Thus, more attributes are considered to calculate the similarity between cases, and consequently the time to find the most similar case is incremented. On the other hand, we have also analyzed a set of bigger sized KBs than for the DT network, which also increases the computing time. However, considering a KB of 50,000 cases, R-CBR employs 110 ms to classify a lightpath (i.e., more than one order of magnitude faster than the Q-Tool). If FixE-CBR is employed, the computing time is reduced to 32 ms as the KB is also reduced to 12,633 cases. Therefore, the cognitive estimator, when using the optimized KB, is approximately two orders of magnitude faster than the Q-Tool.

As we have previously discussed, the cognitive QoT estimator relies on a hybrid approach in order to classify lightpaths; first deciding by means of a threshold length, and then, if required, by a CBR system. That hybrid approach significantly reduces the computing time. However, it should be noticed that even if the cognitive QoT estimator relied exclusively on the CBR system, it would be still faster than the Q-Tool. For instance, considering the DT network equipped with 32 wavelengths per link, the computing time is 6.6 ms (per lightpath) when the CBR system is used versus 768.7 ms (per lightpath) when the Q-Tool is used instead (due to the numerical simulations that take place therein). On the other hand, even if the hardware version of the Q-Tool were used (which reduces the computing time $\sim$28 times when compared with the MATLAB implementation of the Q-Tool [6]), the MATLAB-based cognitive QoT estimator would still be faster.

B. Deployment of the Underlying Knowledge Base

A key issue on the utilization of the cognitive QoT estimator is the deployment of the underlying KB. In this subsection we analyze two pragmatic methods to fill that KB before starting the dynamic operation of the network.

A first option consists in running a set of off-line physical layer simulations emulating different configurations of the network and recording the QoT evaluation of the different light paths. Since this may be a slow and tedious procedure, and thus not too many cases can be compiled, we have analyzed the performance of the estimator when the KB consists of a reduced number of cases (but representing very diverse scenarios). Thus, Fig. 7 shows the percentage of successful classifications of lightpaths for the DT network for a non-optimized KB populated with a low number of cases ($<$500), and also for the optimized version of those KBs (by applying the CRR procedure described in Section II-B). As it can be noticed, even for a non-optimized small KB of 50 cases the percentage of successful clasifications is higher than 98.7%.

A second option consists in filling the KB by gathering experimental data from the optical network, prior to its dynamic operation. For that aim, the network operator may use the network management system to automatically setup and test a reduced number of pragmatic configurations corresponding to different network loads (i.e., to different numbers of lightpaths established, and thus corresponding to different scenarios in terms of co-propagating lightpaths) that are expected to be faced by the network. Therefore, for each configuration, a set of lightpaths are established in the network, and their QoTs are measured by means of network monitors. That information is then used to create the initial KB.

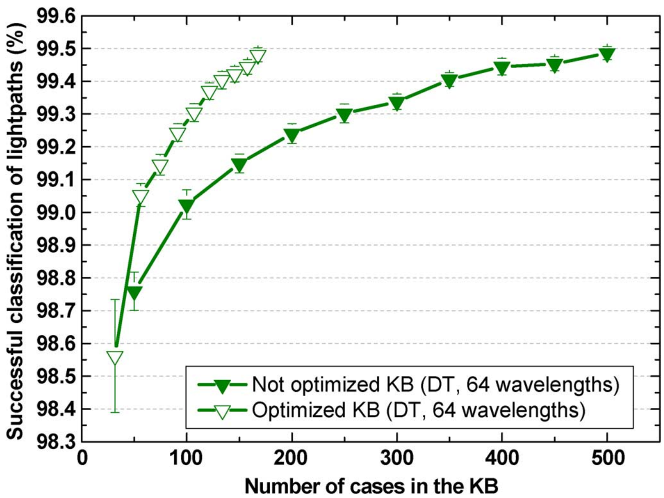

Fig. 7. Successful classifications of lightpaths for small KB sizes (<500 cases) with and without optimization of the KB for DT network and 64 wavelengths.

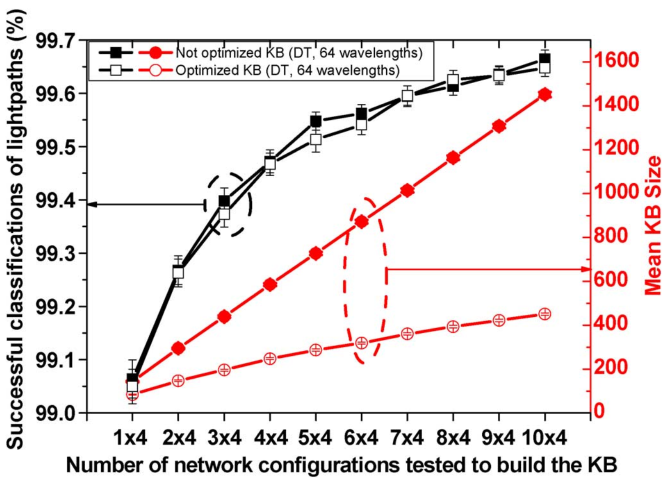

Fig. 8. Successful classifications of lightpaths and KB size for different number of networks configurations tested to build the KB with and without optimization of the KB for DT network with 64 wavelengths.

Fig. 8 shows the simulation results obtained for the DT network, equipped with 64 wavelengths, when using this procedure to build the KB.

In order to build the KB, we have set four random configurations, each representative of a different traffic load: low (0.5), medium (1.5), high (2.5), and very high (4.1). Since we are using a simulation environment and do not have real monitors, we have replaced the measurements of the monitors by the values provided by the Q-Tool when assessing those four configurations.

Then, the performance of the cognitive QoT estimator with that underlying KB has been evaluated for varying network loads ranging from the lowest (0.5) to the highest one (4.1), but also including intermediate traffic loads not considered when building the KB. In order to obtain statistically meaningful results, this procedure has been repeated for 100 different underlying KBs built as previously described.

On the other hand, we have also built bigger KBs by gathering the results of more than one random configuration for each of the four traffic loads (thus, labeled in the x-axis of Fig. 8 as 2×4, 3×4, and so on), and the performance of the QoT estimator has again been analyzed.

Thus, Fig. 8 represents the percentage of successful classifications as a function of the number of network configurations tested to build the KB. Moreover, the resultant size of the KB is also represented on the same figure. There are two families of results: those obtained when using non-optimized KBs (filled symbols), and those obtained when using KBs optimized off-line by means of the CRR technique (empty symbols).

As it can be observed, the percentage of successful classifications reaches up to 99.65% when testing $10 \times 4 = 40$ network configurations to build the KB (leading to a KB of ~1450 cases—excluding those outside of the uncertainty area—). In that scenario, if an optimized KB is used, the success ratio keeps the same but the KB size is significantly reduced (~450 cases). Moreover, by only using the information gathered when testing 4 network configurations (only one configuration for each of the four traffic loads), the underlying KB consists of only 84 cases and the success ratio is higher than 99%.

IV. Conclusion

We have shown that cognition can be successfully applied in optical networks for quality of transmission assessment. In particular, we have proposed a novel cognitive QoT estimator (R-CBR), which is able to quickly determine whether lightpaths comply with quality of transmission requirements in wavelength-routed optical networks. It achieves a high percentage of successful assessments (>99% when applied to DT network and ~98% when applied to the GÉANT2), and moreover, it is much faster (close to three orders of magnitude for the DT network and more than one order of magnitude for the GÉANT2 network) than a past non-cognitive proposal, the Q-Tool [3], [4]. Besides presenting the fundamentals of the estimator, we have also shown how its performance can be improved by incorporating learning and forgetting strategies with the aim of optimizing the underlying KB. Thus, we have introduced a new technique, FixE-CBR, which gets similar or even slightly better success ratios when compared with the use of R-CBR, but with a significant reduction in the number of cases stored in the underlying KB, which in turn translates in a reduction of the computing time. In this way, FixE-CBR is approximately one order of magnitude faster (for on-line operation) than R-CBR for both networks, and approximately four and two orders of magnitude faster than the Q-Tool for the DT and GÉANT networks, respectively.

Although we have focused on the off-line optimization of the KB, the learning and forgetting technique that we have introduced could be used so that the cognitive QoT estimator was able to adapt itself, on real time, to a changing environment. In this way, the underlying KB would evolve to reflect network changes, such as component aging or deterioration.

Moreover, the system has been tested under two different networks, the DT network and the pan-European GÉANT2 network in order to address scalability issues. In this way, it has been proved that both QoT estimators (R-CBR and FixE-CBR) performs slightly worse in GÉANT2 network due to the different distribution of cases, even when incrementing the KB size. However, the percentage of successful classifications is higher than 99% for the 64 wavelengths scenario.

On the other hand, we have described two pragmatic methods to populate the underlying KB of the cognitive QoT estimator. In particular, we have shown that by testing a reduced number of network configurations before starting its dynamic operation, enough information can be gathered to build the KB and achieve a high success ratio (>99% in the DT network).

Finally, it is worthy to note that, although in this paper we have analyzed 10 Gb/s OOK networks for comparison purposes, the case-based reasoning technique proposed is flexible and generic enough to be used in other networking scenarios. As an example, we have recently demonstrated its application in a WDM 80 Gb/s PDM-QPSK dispersion-compensated testbed [10], thereby showing its potential for higher transmission rate systems.

References

[1] S. Azodolmoky *et al.*, "A survey on physical layer impairments aware routing and wavelength assignment algorithms in optical networks," *Comput. Netw.*, vol. 53, no. 7, pp. 926–944, May 2009.

[2] C. V. Saradhi and S. Subramaniam, "Physical layer impairment aware routing (PLIAR) in WDM optical networks: Issues and challenges," *IEEE Commun. Surveys Tutorials*, vol. 11, no. 4, pp. 109–130, Dec. 2009.

[3] S. Azodolmoky *et al.*, "Experimental demonstration of an impairment aware network planning an operation tool for transparent/translucent optical networks," *J. Lightw. Technol.*, vol. 29, no. 4, pp. 439–448, 2011.

[4] S. Azodolmolky, P. Kokkinos, M. Angelou, E. Varvarigos, and I. Tomkos, "DICONET NPOT: An impairments aware tool for planning and managing dynamic optical networks," *J. Netw. Syst. Manag.*, vol. 20, no. 1, pp. 116–133, 2012.

[5] R. Schmogrow *et al.*, "Error vector magnitude as a performance measure for advanced modulation formats," *IEEE Photon. Technol. Lett.*, vol. 24, no. 1, pp. 61–63, Jan. 2012.

[6] Y. Qin *et al.*, "Demonstration of C/S based hardware accelerated QoT estimation tool in dynamic impairment-aware optical network," presented at the Proc. ECOC 2010, 2010, Paper P5.17.

[7] N. Fernández *et al.*, "Cognition to design energetically efficient and impairment aware virtual topologies for optical networks," in *Proc. ONDM 2012*, April 2012.

[8] P. Poggiolini, "The GN model of non-linear propagation in uncompensated coherent optical systems," *J. Lightw. Technol.*, to be published.

[9] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *Artif. Intell. Commun.*, vol. 7, no. 1, pp. 39–59, 1994.

[10] A. Caballero *et al.*, "Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems," *Opt. Exp.*, vol. 20, no. 26, pp. B64–B70, December 10, 2012.

[11] T. Jiménez *et al.*, "A cognitive system for fast quality of transmission estimation in core optical network," presented at the OFC/NFOEC 2012, Los Ángeles, CA, USA, March 2012, Paper OW3A.5.

[12] T. Jiménez *et al.*, "Optimization of the knowledge base of a cognitive quality of transmission estimator for core optical networks," in *Proc. 16th Optical Network Design and Modeling (ONDM 2012)*, Colchester, U.K., Apr. 2012.

[13] CHRON Project. Deliverable D2.1, Specification of the Network Scenario and Service Requirements Nov. 2011 [Online]. Available: http://www.ict-chron.eu

[14] R. López *et al.*, "Retrieval, reuse, revision and retention in case-based reasoning," *The Knowledge Engineering Review*, vol. 30, no. 3, pp. 215–240, 2006.

[15] M. Angelou, Y. Pointurier, D. Careglio, S. Spadaro, and I. Tomkos, "Optimized monitor placement for accurate QoT assessment in core optical networks," *J. Lightw. Technol.*, vol. 4, no. 1, pp. 15–24, Jan. 2012.

[16] D. W. Aha, "Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms," *Int. J. Man-Machine Studies*, vol. 36, pp. 267–287, 1992.

[17] D. Wettscherek, D. W. Aha, and T. Mohri, "A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms," *Artif. Intell. Rev.*, vol. 11, no. 1–5, pp. 273–314, Feb. 1997.

[18] M. Salamó and M. López-Sánchez, "Adaptive case-base reasoning using retention and forgetting strategies," *Knowledge-Based Syst.*, vol. 24, no. 2, pp. 230–247, Mar. 2011.

[19] S. J. Delany and P. Cunningham, "An analysis of case-base editing in a spam filtering system," in *Advances in Case-Based Reasoning*. New York: Springer, 2004, pp. 3–25.

[20] E. McKenna and B. Smyth, "Competence-guided editing methods for lazy learning," in *Proc. 14th Eur. Conf. Artif. Intell.*, 2000, pp. 60–64.

[21] A. Mokhtar and M. Azizoglu, "Adaptive wavelength routing in all-optical networks," *IEEE/ACM Trans. Networking*, vol. 6, no. 2, pp. 197–206, April 1998.

[22] A. Zapata-Beghelli and P. Bayvel, "Dynamic versus static wavelength-routed optical networks," *J. Lightw. Technol.*, vol. 26, no. 10, pp. 3403–3415, 2008.

Tamara Jiménez received the Telecommunication Engineer degree from the University of Valladolid, Spain, in 2008, where she is currently working toward the Ph.D. degree in the Optical Communications Group.

Her current research focuses on the design and performance evaluation of optical networks (especially long reach passive optical networks, and cognitive optical networks).

Juan Carlos Aguado received the Telecommunication Engineer and Ph.D. degrees from the University of Valladolid, Spain, in 1997 and 2005, respectively.

He has worked as a Junior Lecturer at the University of Valladolid since 1998. His current research interests include the design and evaluation of cognitive methods applied to physical-layer modeling and traffic routing in heterogeneous optical networks.

Ignacio de Miguel received the Telecommunication Engineer degree in 1997, and the Ph.D. degree in 2002, both from the University of Valladolid, Spain.

Since 1997 he has worked as a Lecturer at the University of Valladolid. He has also been a Visiting Research Fellow at University College London (UCL), working in the Optical Networks Group. His research interests are the design and performance evaluation of optical networks, especially hybrid optical networks, cognitive optical networks, as well as IP over WDM.

Dr. de Miguel is the recipient of the 1997 Innovation and Development Regional Prize for his Graduation Project, and the Nortel Networks Prize to the best Ph.D. Thesis on Optical Internet in 2002, awarded by the Spanish Institute and Association of Telecommunication Engineers (COIT/AEIT).

Ramón J. Durán was born in Cáceres, Spain, in 1978. He received the Telecommunication Engineer degree in 2002 and the Ph.D. degree in 2008, both from the University of Valladolid, Spain.

Since 2002, he has worked as a Junior Lecturer at the University of Valladolid and currently he is also Deputy Director of the Faculty of Telecommunication Engineering. His current research focuses on the design and performance evaluation of cognitive heterogeneous optical networks. Dr. Durán is the author of more than 60 papers in international journals and conferences.

Marianna Angelou (S'09–M'12) received the Ph.D. degree from Universitat Politècnica de Catalunya, Barcelona in April 2012.

In January 2008, she joined the "High-speed Networks and Optical Communications" Group of Athens Information Technology, as a Research Scientist. She currently holds the position of Research Project Manager of EC-funded collaborative projects on behalf of Optronics Technologies SA, Athens. Her research activities focus on cross-layer optimization techniques for optical networks and cover a broad range of topics in that area including physical layer modeling, energy-efficiency, and networking with flexible/adaptive transmission characteristics.

Noemí Merayo received the Telecommunication Engineer degree in engineering from the Valladolid University, Spain, in February of 2004 and the Ph.D. degree in the Optical Communication Group at the University of Valladolid, in July 2009.

Since 2005 she has worked as a Junior Lecturer at the University of Valladolid. He has also been a Visiting Research Fellow at the University of Hertfordshire, London, U.K., working in the Optical Networks Group, Science and Technology Research Institute (STRI). Recently, she has been a postdoctoral research in the group of "Transmisiones Ópticas de Banda Ancha (TOyBA)" of the University of Zaragoza. Her current research focuses on the design and performance evaluation of optical networks, especially passive optical networks.

Patricia Fernández received the Telecommunication Engineer degree from Universidad Politécnica de Cataluña, Barcelona, Spain, in 1997 and the Ph.D. degree in 2004 from University of Valladolid.

Since 1999, she has worked as a Junior Lecturer at the University of Valladolid. Her research interests are passive optical networks and fiber-optic communications components. Dr. Fernández is the author of more than 40 papers in international journals and conferences.

Rubén M. Lorenzo received the Telecommunication Engineer and Ph.D. degrees from the University of Valladolid, Spain, in 1996 and 1999, respectively. From 1996 to 2000, he was a Junior Lecturer at the University of Valladolid, and joined the Optical Communications Group.

Since 2000, he has been a Lecturer. His research interests include Integrated Optics, Optical Communication Systems and Optical Networks. He is currently the Head of the Faculty of Telecommunication Engineering at University of Valladolid.

Ioannis Tomkos has been with AIT since September 2002 (serving as Professor, Research Group Head and Associate Dean).

In the past he was Senior Scientist at Corning Inc., USA (1999–2002) and Research Fellow at University of Athens, Athens, Greece (1995–1999). He has represented AIT as Principal Investigator in about 20 European Union and industry funded research projects (including 9 currently active projects) and has a consortium-wide initiator/leader role. His fields of expertise are telecommunication systems, networks and photonics, as well as techno-economic analysis and business planning of ICTs. Together with his colleagues and students he has authored about 450 peer-reviewed archival scientific articles, including over 120 journal/magazine/book publications and 330 conference/workshop proceedings papers.

Dr. Tomkos was elected in 2007 as Distinguished Lecturer of the IEEE Communications Society for the topic of optical networking. He has served as the Chair of the International Optical Networking Technical Committee of the IEEE Communications Society (2007–2008) and the Chairman of the IFIP working group on Photonic Networking (2008–2009). He is currently the Chairman of the OSA Technical Group on Optical Communications (2009–2012) and the Chairman of the IEEE Photonics Society Greek Chapter (2010–2012). He is also Chairman of the working group "Next Generation Networks" of the Digital Greece 2020 Forum. He has also been General Chair, Technical Program Chair, Subcommittee Chair, Symposium Chair, or/and member of the steering/organizing committees for the major conferences in the area of telecommunications/networking (more than 100 conferences/workshops). In addition, he is/was a member of the Editorial Boards of the IEEE/OSA JOURNAL OF LIGHTWAVE TECHNOLOGY, the IEEE/OSA JOURNAL OF OPTICAL COMMUNICATIONS AND NETWORKING, the IET JOURNAL ON OPTOELECTRONICS, and the International Journal on Telecommunications Management.

Evaristo J. Abril received the Telecommunication Engineer and Ph.D. degrees from the Politécnica de Madrid, Spain, in 1985 and 1987, respectively.

From 1984 to 1986, he was a Research Assistant at Universidad Politécnica de Madrid, becoming Lecturer in 1987. Since 1995, he has been Full Professor at University of Valladolid, Spain, where he founded the Optical Communications Group. His research interests include Integrated Optics, Optical Communication Systems and Optical Networks.

Prof. Abril is the author of more than 100 papers in international journals and conferences.