A SVM Approach for Lightpath QoT Estimation in Optical Transport Networks

Javier Mata, Ignacio de Miguel, Ramón J. Durán, Juan Carlos Aguado, Noemí Merayo, Lidia Ruiz, Patricia Fernández, Rubén M. Lorenzo, Evaristo J. Abril

Optical Communications Group, Department of Signal Theory and Communications and Telematics Engineering
Universidad de Valladolid (UVa)
Valladolid, Spain
E-mail: fcojavier.mata@alumnos.uva.es, ignacio.miguel@tel.uva.es, rduran@tel.uva.es

Abstract—**A novel quality of transmission (QoT) estimator based on support vector machines (SVM) is proposed for classifying optical connections (lightpaths) into high or low quality categories in impairment-aware wavelength-routed optical networks (WRONs). The performance of the SVM-based estimator is evaluated in a long haul communications network and compared to previous semi-analytical and cognitive proposals. Results show that the SVM approach significantly reduces the necessary computing time to estimate the QoT of a given lightpath, critical aspect of design in these networks, and even slightly improves accuracy.**

Keywords-Case-Based Reasoning (CBR), impairment-aware networking, lightpath, quality of transmission (QoT), Support Vector Machines (SVM), Wavelength-Routed Optical Networks (WRON)

I. INTRODUCTION

Optical networks performing wavelength division multiplexing (WDM) are made up of *optical fibers*, which can transport different data channels simultaneously by using different wavelengths (frequencies), and *optical nodes*, which can perform optical routing based on wavelengths. Wavelength-routed optical networks (WRONs) are WDM-based all-optical networks (i.e., incoming wavelengths are switched to the output fibers optically, with no conversion to the electrical domain) defined by lightpaths or optical circuit-switched connections between two network nodes, not necessarily adjacent, characterized by a given route-wavelength pair. The problem of selecting an appropriate route and wavelength among the many possible alternatives for each connection so that two routes sharing a link are not assigned the same wavelength is named the Routing and Wavelength Assignment (RWA) problem.

However, the classical RWA problem does not take into account the physical impairments that may degrade optical signals as they propagate through active and passive optical components towards their destination. Since, as mentioned above, in transparent networks no regeneration is performed in intermediate nodes, these physical impairments will affect the quality of transmission (QoT) of the signal, which may not comply with service requirements. Therefore, impairment-aware (IA) networking needs to be incorporated to the classical RWA problem (thus, IA-RWA, from now on)

so as to find a route and an available wavelength to establish a lightpath in such a way that enough quality for data transmission is ensured. For this reason, lightpaths QoT estimation prior to deployment becomes of paramount importance. Moreover, since in dynamic WRONs lightpath requests arrive at the Path Computation Element (PCE, a centralized entity in charge of solving the IA-RWA problem) in real time, and the PCE must solve this IA-RWA problem within a very short period of time, not only the accuracy of the QoT estimation of the lightpaths is crucial, but also the computing time required to get this estimate, which is obtained by means of a model previously built.

Cognitive approaches have been proposed to address this issue, like in [1], where a QoT estimator provides with a highly accurate classification of lightpaths into high and low categories by performing Case-Based Reasoning (CBR) [2] in a 10 Gb/s system. This approach, several magnitude orders faster than previous semi-analytical ones (e.g. Q-Tool [3]), has been indeed experimentally demonstrated in more advanced systems operating up to 80 Gb/s [4]. One step beyond, the incorporation of learning and forgetting techniques in order to reduce the size of the knowledge base (KB) on which CBR relies, thus decreasing computing time, is presented in [5]. So far, few studies have applied machine learning techniques to solve this particular problem. As an example, [6] makes use of random forests. However, this study is carried out on synthetic data and without yielding results in computing time, critical aspect of design for practical implementations.

We now propose to replace the CBR module of the QoT estimator presented in [5] with an element based on SVM in order to take advantage of the effectiveness of this technique in binary classification problems. This novel QoT estimator, trained with a set of established lightpaths obtained by previous off-line simulations over the network under study, significantly reduces the necessary computing time to estimate the QoT of a given lightpath in comparison with previously described proposals and even slightly improves the accuracy.

The remainder of the paper is structured as follows: Section II describes the novel QoT estimator based on SVM, while Section III discusses simulation scenarios and results and Section IV gathers the main conclusions of the study.

II. DESCRIPTION OF THE QoT ESTIMATOR

We have developed a QoT estimator that classifies lightpaths into high and low quality categories based on a user-defined Q-factor threshold, an indicator of the quality of transmission which is inversely related to the bit error rate (BER) [7]. Specifically, this threshold is set to 16.9 dB, which corresponds to a BER level of 10^{-12}. Thus, lightpaths labeled with a Q-factor value above the threshold, i.e., with a tolerable (low) BER, are considered to comply with quality requirements (high quality category) and those below the threshold are considered non-viable lightpaths.

Among the numerous factors that may influence on the QoT of a lightpath, its length is probably the one with the greatest impact. This is why, as conducted in [5], the proposed estimator is of a hybrid nature, i.e., when the time comes to classifying a lighpath, its length is taken into account in first place and, only if this variable is not sufficiently determinant, the SVM-based module becomes the one in charge of the prediction. In the end, the problem distinguishes three different areas according to the length of the lightpath. A first area is defined up to a certain length limit, below which all lightpaths can be considered high quality. In contrast, a second area in which every lighpath is considered low quality, is defined above a certain length limit. The limits of these two areas are determined so that the probability of successful classification of lightpaths within them is higher than 99.99%. Finally, between both limits, a third area, in which the rest of the characteristics of the lighpaths also needs to be considered as they play an important role in calculating the Q-factor, is specified. The latter, so called uncertainty area, is the one solved by the SVM-based module, while the rest are managed by a simple decision maker.

The goal of SVM is to learn a mapping function between the input features (i.e., the observed variables) and the target value (i.e., the class labels) [8]. For this purpose, SVM needs to be provided with a data set from which training, validation and testing cases can be extracted in order to build the model and study its performance. Particularly, in this case, the starting data set is obtained by running previous off-line simulations of the IA-RWA problem over the network under study. The cases thus collected, i.e., the different lightpaths established at different moments and with different network traffic loads, consist of a set of attributes that describe the lightpath and their associated Q-factors. Specifically, the description of the lightpath contains the source and destination nodes, the set of links it traverses (represented by the percentage of its individual contribution to the total length of the lightpath), the corresponding wavelength, its total length, the sum of co-propagating lightpaths per link and the standard deviation of that number. In addition, the associated Q-factor is calculated also off-line by means of the Q-Tool [3], accurate yet too slow.

Since the relationship between the QoT categories and the lightpaths' attributes is nonlinear, a radial basis function kernel (RBF) is considered to nonlinearly map samples into a higher dimensional space [8]. Initially a training phase is carried out, where different data subsets belonging to the uncertainty area are provided to the SVM module in order to learn the corresponding mapping function between the attributes and the quality categories. A 10-fold cross validation process is then performed to characterize thoroughly the expected performance of the model thus constructed. This validation also includes an optimization process to look for the tuning of parameters that offers the best accuracy in classification, which, considering a RBF kernel, consists on finding the best combination of C, or penalty parameter and γ, the kernel parameter. Ranges from 2^{-5} to 2^{15} and 2^{-15} to 2^{3} are considered for C and γ, respectively [8].

Once the model is built and characterized, the performance of the system is tested with different new data subsets. Instances from these subsets are divided into cases belonging to the uncertainty area, which are classified by the SVM module and cases outside the uncertainty area, which are classified by means of a simple decision maker as stated before. As output variables of this phase, the accuracy in the classification of lightpaths carried out by the joint action of the decision maker and the SVM module and the computing time to classify a lightpath are obtained.

III. SIMULATION SCENARIOS AND RESULTS

To evaluate the performance of the proposed QoT estimator, simulations have been carried out in a long haul network, the 14-node Deutsche Telekom (DT) network [3], considering 32 and 64 wavelengths per link. For the sake of a fair comparison, the data set, i.e., the set of established lightpaths obtained by means of offline simulations solving the IA-RWA problem over the DT network, is the same as in [5]. In addition, our novel QoT estimator is also implemented in MATLAB and simulations are ran in the same computer (Debian GNU/Linux 6.0 machine using one core of an AMD Opteron 6128 processor) and in the same exact conditions as [5], which includes the same length limits of the uncertainty area that are shown in Table I, the same number of training/testing cases and the same number of repetitions of the experiment.

Specifically, our QoT estimator has been initially trained (and cross-validated) with different number of randomly chosen cases belonging to the uncertainty area, running from 500 to 5000. Once the training/cross-validation phase was finished, the performance of the system was analyzed (test phase). For that aim, other 6000 cases, i.e., lightpaths, belonging to both the certainty and uncertainty areas have been evaluated.

The results presented in the following figures have been obtained after repeating this process 100 times with different data subsets. Average results are presented together with 95% confidence intervals, although most of the times the confidence intervals are smaller than the size of the symbols.

TABLE I.   LOW AND HIGH LIMITS OF THE UNCERTAINTY AREA

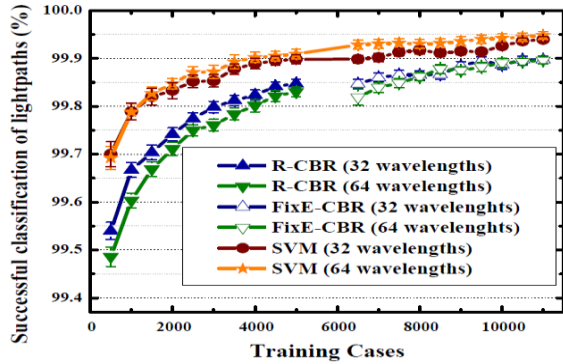| Network	Number of wavelengths	Low length limit (km)	High length limit (km)
DT	32	975	1875
	64	975	2050

Figure 1. Successful classification of lightpaths for R-CBR, FixE-CBR and SVM approaches over DT.
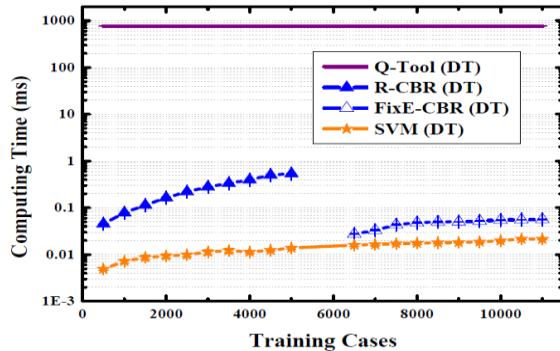

Figure 2. Computing time per lightpath for Q-Tool, R-CBR, FixE-CBR and SVM approaches over DT.

As the study conducted by Jiménez et al. [5] has been used as benchmark to evaluate the performance of the novel proposal, it is worth explaining briefly the two approaches based on CBR that were used in such study: R-CBR performed estimations relying on a static KB, while FixE-CBR included an additional optimization phase of the KB by applying learning and forgetting techniques based on an extra set of 6000 lightpaths. This optimization aimed at reducing the size of the KB on which CBR relied, thus decreasing computing time. Although the final size of the KB was indeed much smaller than in R-CBR, in the end, 6000 extra cases were used in a pseudo-training phase.

Fig. 1, shows the percentage of scuccessful classification of lightpaths into high and low QoT categories for both the CBR-based and the SVM-based approaches, over DT network. As expected, increments on the number of training cases imply, in turn, improvements of the performance in terms of accuracy in classification for every proposal. When comparing the results among the different approaches, SVM offers the highest accuracy in classification considering the same size of the training data set, reaching up to a 99.95% percentage of success for a training data set of 11000 cases.

As we have mentioned before, not only the accuracy in classification is important but also the computing time to classify a lightpath is crucial, especially in highly dynamic networks. Fig. 2, presents the comparison in computing time per lightpath for the different approaches previously

mentioned. Q-Tool, based on semi-analytical models, may be accurate but it is by far the slowest one. The significant reductions in computing time per lightpath offered by both CBR approaches with respect to Q-Tool are significantly outperformed by our proposal based on SVM. As a matter of fact, SVM is up to almost two magnitude orders faster than R-CBR and more than three times faster than the optimized CBR (FixE-CBR). Moreover, since SVM classifies lightpaths using a mapping function and thus the computing time remains almost constant through the training dataset sizes' range, the computing time corresponding to the highest accuracy (11000 cases) is lower than every other model regardless the point analysed.

IV. CONCLUSIONS

Accurate and fast lightpath QoT estimation is of paramount importance when solving the IA-RWA problem in optical networks. We have proposed a novel QoT estimator based on SVM, which has proven to be efficient in solving this particular problem outperforming previous analytical and cognitive CBR-based proposals both in accuracy, reaching up to a 99.95% success in classification of lightpaths, and in computing time, improving R-CBR approach in almost two magnitude orders and classifying lightpaths more than three times faster than FixE-CBR.

REFERENCES

[1] T. Jiménez, et al., "A cognitive system for fast quality of transmission estimation in core optical networks," in Proc. OFC/NFOEC 2012.

[2] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," AI communications, vol. 7, no. 1, pp. 39–59, 1994.

[3] S. Azodolmolky, et al., "Experimental demonstration of an impairment aware network planning and operation tool for transparent/translucent optical networks," Journal of Lightwave Technology, vol. 29, no. 4, pp. 439–448, 2011.

[4] A. Caballero, et al., "Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems," Optics Express, vol. 20, no. 26, pp. B64–B70, 2012.

[5] T. Jiménez, et al., "A cognitive quality of transmission estimator for core optical networks," Journal of Lightwave Technology, vol. 31, no. 6, pp. 942–951, 2013.

[6] L. Barletta, A. Giusti, C. Rottondi, and M. Tornatore, "QoT estimation for unestablished lighpaths using machine learning," in Proc. OFC 2017.

[7] R. Schmogrow et al., "Error Vector Magnitude as a Performance Measure for Advanced Modulation Formats," in IEEE Photonics Technology Letters, vol. 24, no. 1, pp. 61-63, Jan.1, 2012.

[8] C. W. Hsu, C. C. Chang, and C. J. Lin, "A Practical Guide to Support Vector Classification," 2008. [Online]. Available: http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.