A BiLSTM-CNN based Multitask Learning Approach for Fiber Fault Diagnosis

Khouloud Abdelli1,2, Helmut Grießer1, Carsten Tropschug2, and Stephan Pachnicke3

1ADVA Optical Networking SE, Fraunhoferstr. 9a, 82152 Munich/Martinsried, Germany
2ADVA Optical Networking SE, Märzenquelle 1-3, 98617 Meiningen, Germany
3Christian-Albrechts-Universität zu Kiel, Kaiserallee 2, 24143 Kiel, Germany
E-mail: KAbdelli@adva.com

Abstract: A novel multitask learning approach based on stacked bidirectional long short-term memory (BiLSTM) networks and convolutional neural networks (CNN) for detecting, localizing, characterizing, and identifying fiber faults is proposed. It outperforms conventionally employed techniques. © 2021 The Authors

1. Introduction

Optical time reflectometry (OTDR), a technique based on Rayleigh backscattering, has been widely applied for fiber characteristics’ measurements and for fiber fault detection and localization. However, OTDR traces can be hard to interpret even by experienced field crews. Fiber fault real-time detection is considered as the industry standard to ensure optical network survivability and reliability. Therefore, it is crucial to develop an automated reliable technique capable of accurately and quickly detecting, localizing, and identifying faults given noisy OTDR data and thereby reducing operation-and-maintenance expenses (OPEX). In the past several OTDR event analysis approaches have been proposed. Conventionally, the OTDR event detection technique relies on a two-point method combined with least square approximation, which is coarse and noise sensitive. Recently, data-driven approaches based on Long Short-Term Memory (LSTM) networks and convolutional neural networks (CNN) for detecting, locating, characterizing, and identifying fiber faults have been proven to efficiently solve fault diagnosis problems given sequential data [1-2] and could provide an untapped potential to tackle the fiber fault diagnosis problem.

In this paper, we develop a novel hybrid multitask learning framework combining bidirectional long short-term memory (BiLSTM) networks and convolutional neural networks (CNN) to diagnose optical fiber faults when analyzing noisy OTDR data. The multi-task learning architecture is composed of a BiLSTM-CNN shared hidden layer distributing the knowledge across multiple tasks (fault detection $T_1$, fault localization $T_2$, fault characterization $T_3$ and fault cause identification $T_4$) followed by a specific task layer. The proposed approach is applied to noisy experimental OTDR data, whose SNR (signal-to-noise ratio) values vary from 0 dB to 30 dB, comprising multiple fiber faults with various characteristics. The experimental results show that our model (i) achieves better fault diagnosis capability compared to several other ML techniques namely CNN, LSTM, and BiLSTM and (ii) outperforms the conventional OTDR event analysis methods particularly for low SNR values.

2. Setup & Configurations

The SNR of the OTDR trace depends on the receiver photodiode sensitivity, the pulse power used in the measurement, the attenuation, the fiber length, and the averaging of repeated measurements or low-pass filtering of the trace. As the fiber is given and receiver sensitivity as well as launch power are limited, the only parameters left to influence the SNR, are the pulse width, the filtering (which allow to trade SNR against location accuracy) and the averaging (which can improve the SNR at the cost of a longer measurement time). Therefore, for our experiments, the SNR is considered as the relevant figure-of-merit and averaging is used to influence it.

2.1. Experimental Setups

The experimental setups shown in Fig. 1 are conducted for recording OTDR traces incorporating different types of fiber events with various reflectances and losses. Optical components like connectors, VOAs, reflectors are also used to model specific faults in the fiber link. To generate non-reflective faults, involving only attenuation and no reflection, with small decrease, an angled physical contact (APC) connector is placed at the end of the 1 km fiber depicted in Fig. 1 (left). These non-reflective faults are identified as fiber-bend. To produce non-reflective events with abrupt decrease, an APC connector is located at the end of setup 1 (i.e. the end of the 5 km fiber). The generated non-reflective faults are classified in that case as tilted fiber-cut. The reflective faults (Fresnel reflection) with small peaks are induced by placing a reflector at the end of the 1 km fiber shown in Fig. 1 (right). They are identified as longitudinal connector-misalignment. The reflectance of these events is varied. By putting a physical contact (PC) connector at the end of the setup 1, reflective events with sharp peak and abrupt decrease are caused. They are classified in such circumstances as perpendicular fiber-cut. The coupler, the variable optical attenuators (VOA) and the fibers are individually connected with patch cords. Every connection between the aforementioned components
We use several individual task the proposed model is depicted in transferred pooling learning and to enhance the generalization capability. Given that the tasks are highly related and can significantly benefit from the knowledge sharing across them, a shared hidden layer consists of one BiLSTM layer with 32 cells followed by CNN layers containing mainly one convolutional layer having 32 filters with the max pooling layer succeeded by a dropout layer to prevent overfitting. The knowledge learned by BiLSTM-CNN is then transferred to the four task-specific layers composed of 16, 20, 32, 40 neurons, respectively. The overall structure of the proposed model is depicted in Fig. 3. The total loss of our model is computed as the weighted sum of the four individual task losses set to 1.5, 0.5, 1.8 and 1, respectively.

3. Results and Discussion

We use several metrics including the detection rate (i.e. detection probability) for the $T_1$ evaluation, the root mean...
square error (RMSE) for $T_2$ and $T_3$ performance assessment, and the receiver operating characteristic (ROC) curves summarizing the trade-off between the true positive and false positive rates for different threshold settings for $T_4$ evaluation. As depicted in Fig. 4, the detection probability of the different events at a false alarm rate (FAR) of 0.01 increases with SNR, and for SNR values higher than 10 dB, it is approaching 1. The RMSE of position estimation decreases with SNR, and it is less than 2 m for SNR values higher than 20 dB. The RMSE of the loss ($RMSE_L$) and reflectance ($RMSE_R$) prediction decreases as SNR increases. The $RMSE_L$ is less than 0.3 dB for SNR values higher than 20 dB whereas the $RMSE_R$ can be as low as 2 dB for higher SNR values. The ROC depicted in Fig. 4 shows that our model could distinguish between the different event causes.

We compare the proposed model with different ML methods, i.e. CNN, LSTM and BiLSTM using different metrics namely accuracy, F1 score (a harmonic mean of precision and recall), RMSE, mean absolute error (MAE), symmetric mean absolute percentage error (SMAPE) and the area under curve (AUC). The results shown in Tab. 1 demonstrate that our method outperforms the aforementioned ML techniques in all metrics.

Table 1. Comparison Results of our model with other ML methods

| Metrics                  | Method  | CNN  | LSTM | BiLSTM | Proposed Model |
|--------------------------|---------|------|------|--------|----------------|
| $T_1$: Event detection Accuracy [%] | 89.5 | 91.54 | 91.93 | 92.32 |
| F1 score                 | 0.89    | 0.91 | 0.92 |        |
| $T_2$: Event position estimation RMSE [m] | 3.31 | 3.43 | 2.85 | 2.73 |
| MAE [m]                  | 2.36    | 2.46 | 1.84 | 1.66 |
| $T_3$: Event characterization ($RMSE_L, RMSE_R$) | (10.38, 1.48) | (6.33, 1.03) | (6.15, 1.1) | (4.76, 0.8) |
| ($SMAPE_L, SMAPE_R$)     | (57.43, 60.94) | (54.59, 55.44) | (55.07, 54) | (53.86, 54) |
| $T_4$: Event identification Accuracy [%] | 88 | 89 | 90 | 91 |
| AUC                      | 0.92    | 0.93 | 0.93 | 0.94 |

The proposed framework is compared with a conventional rank-1 matched subspace detector (R1MSDE [4]) using an unseen test dataset including reflective, non-reflective and no event sequences. The results depicted in Fig. 5 show that our approach achieves a notable improvement over the conventional technique particularly for low SNR values.

4. Conclusion

We proposed a multi-task framework based on BiLSTM-CNN for fiber fault diagnosis by processing noisy OTDR data. The experimental results demonstrated that our approach significantly outperforms different ML methods as well as the conventional OTDR event analysis technique.

The work has been partially funded by the German Ministry of Education and Research in the project OptiCON (#16KIS0989K).

5. References

[1] H. Zhao, et al, "Sequential Fault Diagnosis Based on LSTM Neural Network," IEEE Access, 6, 12929-12939, 2018.
[2] K. Abdelli, et al, "Machine Learning based Data Driven Diagnostic and Prognostic Approach for Laser Reliability Enhancement," ICTON, 2020.
[3] W. Lee, et al, "Identification method of non-reflective faults based on index distribution of optical fibers," Opt. Express 22, 325-337 (2014)
[4] F. Liu and C. J. Zarowski, "Detection and location of connection splice events in fiber optics given noisy OTDR data. Part II. R1MSDE method," in IEEE Transactions on Instrumentation and Measurement, vol. 53, no. 2, pp. 546-556, April 2004