Decision Trees for Event Signature Classification on Fiber Optic Cables in Quaternion Coordinates

Essen Dossev(1)(2), Petar Djukic(2), Christine Tremblay(1)

(1) École de Technologie Supérieure, maria.dosseva.1@ens.etsmtl.ca
(2) Ciena Canada, {edosseva, pdjukic}@ciena.com

Abstract *Proximal events posing risks to network service were classified using Decision Trees on State of Polarization Multivariate Time Series data. Aggregate features of interests were individually evaluated to determine their significance, demonstrating that a combination of two aggregates sufficed to produced 98.8% event classification accuracy.*

Introduction

Advancements in digital signal processing with coherent receivers have allowed for detailed monitoring of State of Polarization (SOP) in fiber optic networks. Events occurring near fiber optic cables present themselves as discernible characteristic signatures in SOP measurements. The fiber acts as a medium of propagation for an underlying ground truth which can be used as a sensor for observing surrounding events[1],[2].

Classification of the signatures enables identification of the underlying events. Some of the events previously identified through SOP measurements are fiber break detection[3]–[5], earthquake detection[6], property surveillance[7], and even robotics[8]. One can also imagine other scenarios where SOP is used for surveillance, such as detecting ships or fiber tap detection.

In our work, we explore the use of Decision Trees (DT) for event classification[9]. Implementing ML methods in such applications provides improved automation and less need for hands-on domain-specific knowledge. DT in particular are advantageous for their interpretability.

Multivariate Time Series Data

The dataset is comprised of 16,551 multivariate time series (MTS) SOP recordings, initialized to the North pole of the Poincaré sphere and encoded in the quaternion domain[3] as $a(t), \ldots, d(t)$. The MTS were generated in lab conditions using an Arduino controlled robot arm programmed to produce one of four distinctive event categories: 'bending' (2873), 'shaking' (1922), 'small hit' (6561), and 'up and down' (5195)[1], seen in Figure 1.

[1]Human inspection revealed that the data as originally provided had at least one mislabeled event; the first instance labeled 'small hit' seemed to better fit the pattern of 'shaking'. Numbers presented here incorporate this correction.

The MTS were sampled at a rate of 1920 Hz for a duration of 8128 samples (4.23 seconds). The event triggers had been previously aligned, with 256 samples recorded prior and 7872 samples recorded since the event start.

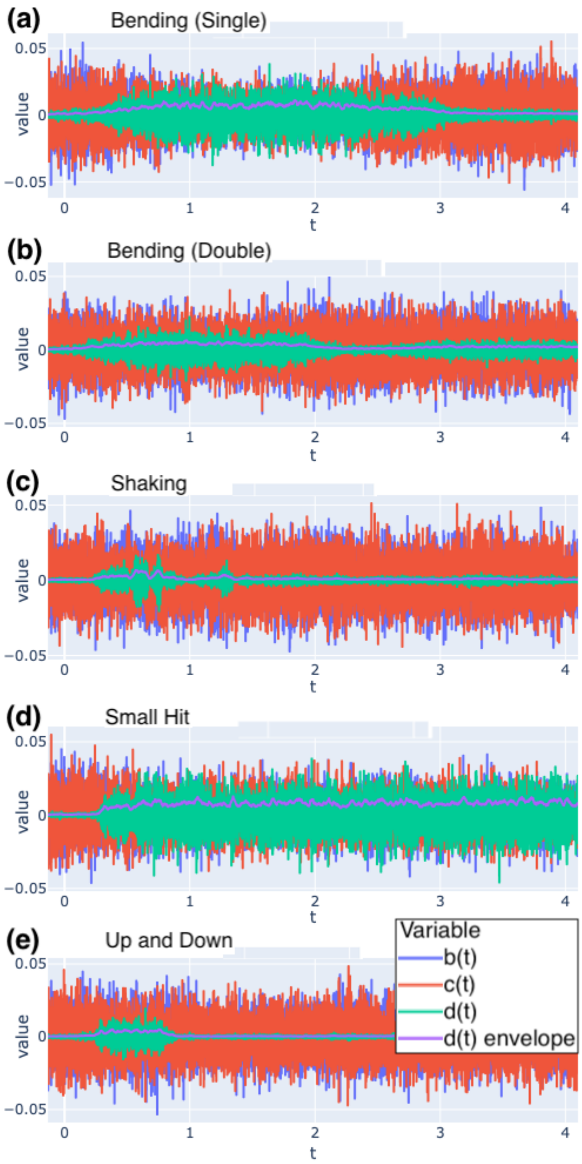

Fig. 1: Samples of the SOP MTS in the quaternion domain.

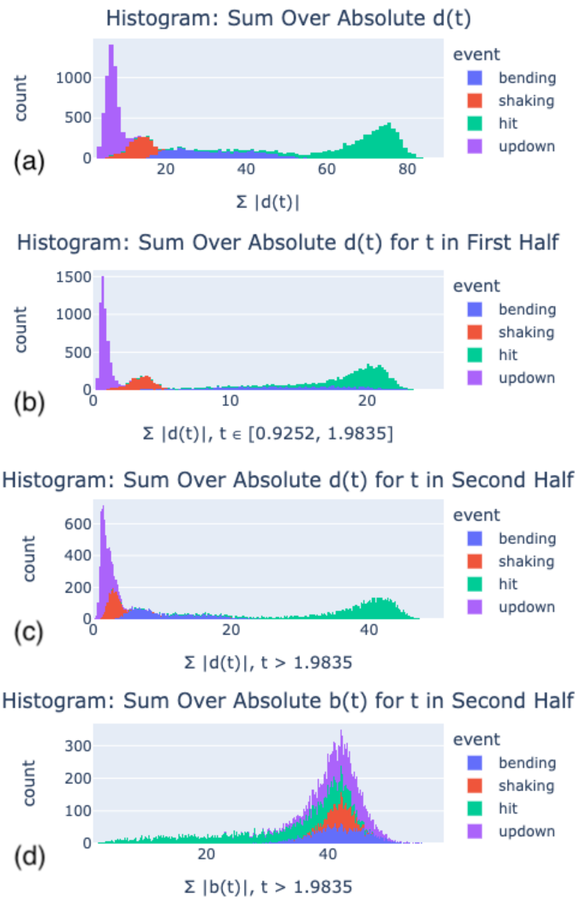

Fig. 2: Histograms over various individual features

Feature Engineering

Aggregate features of interest which had previously been discovered from the dataset[3] were focused on two ranges: $t \in [0.9525, 1.9835]$ and $t > 1.9835$. These correspond roughly to the first half and the second half of the MTS.

In order to aggregate over these two ranges, we exclusively used the summation of absolute values of the variables $b(t)$, $c(t)$ and $d(t)$; we did not use $a(t)$ as it was simply a value of 1, give or take some noise, since the SOP measurements are on the Poincaré sphere; moreover, we did not elect to use other aggregate functions for engineering features, such as mean or max, as these were found to be highly correlated with the sum.

Additionally, the frequency domain was explored for extraction of potential aggregate features, both directly for the MTS as well as for the envelopes of the $d(t)$ variable; however, both were rejected as the plots were less clearly discernible by event type than they had been in the quaternion domain.

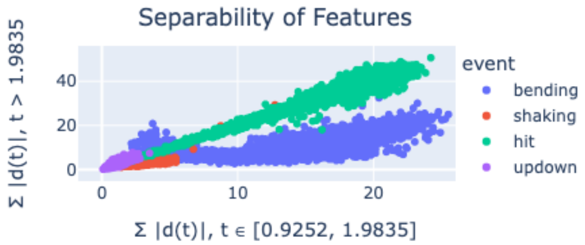

Fig. 3: Separable features: sum absolutes of d(t) over first half and second half

Figure 2 presents the histograms of some features selected for further examination and use in the DT classification; histograms for $b(t)$ in the first half, and $c(t)$ in either the first or second half greatly resembled that of subfigure (d) and were not included due to limited space. Feature scaling was not necessary in preparation for DT.

Many of the features individually do not provide statistical separability of the classes (subfigure (d)); however, when features are examined concurrently they produce a space of separable clusters by event type, as seen in Figure 3.

Anomalous event MTS which clustered within dissimilar event categories revealed the potential that more than one event might appear consecutively within the frame of a single event, as seen in Figure 4. The existence of such events suggests that fixed-size frames for classification are not appropriate, as the length of event signatures varies according to the event type, and may even vary for a single event type, as seen for the 'bending' events in Figure 1.

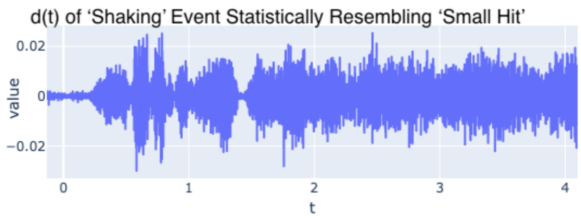

Fig. 4: Anomalous sample: a 'shaking' type event which presents the characteristic shape of of 'small hit' events in the latter half of the frame.

Decision Tree Classification Results

As a baseline classifier, event cluster ranges from the single feature $\sum |d(t)|$ shown in Figure 2(a) were manually evaluated, thus classifying events: 'up and down' (0-10), 'shaking' (10-18), 'bending' (18-48), 'small hit' (48-89). The result of this classification can be seen in Figure 5(a), giving an overall accuracy of 90.9%.

Comparison of aggregate features suggests that those based on the $d(t)$ variable showed that they outperformed both $b(t)$ and $c(t)$ over similar ranges, as seen in Tables 1 and 2; when points

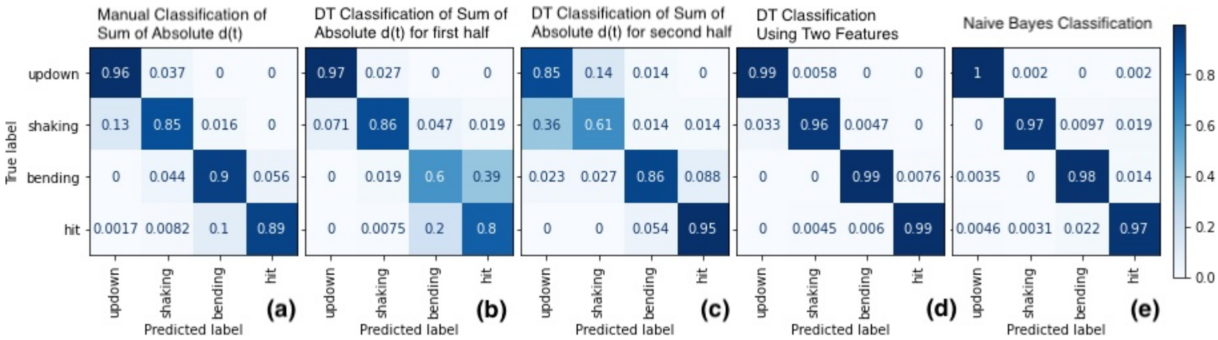

Fig. 5: Confusion matrices

on a sphere cluster about the pole of an axis, the rotations about that axis produce the most identifiable signatures. This can be seen in Figure 1 where the $d(t)$ variable is the most identifiable of the event signatures. In particular, aggregates of $d(t)$ of ranges of either the earlier half or later half of the time series were better at differentiating between 'shaking' and 'up and down' type events, or 'bending' and 'hit' type events, respectively, as seen in Figure 1 (b) and (c).

Tab. 1: Accuracies of single-feature DTs

Feature	Accuracy
$\sum \lvert d(t)\rvert, t \in [0.9252, 1.9835]$	82.8%
$\sum \lvert d(t)\rvert, t > 1.9835$	85.9%
$\sum \lvert b(t)\rvert, t \in [0.9252, 1.9835]$	55.6%
$\sum \lvert b(t)\rvert, t > 1.9835$	54.6%
$\sum \lvert c(t)\rvert, t \in [0.9252, 1.9835]$	54.2%
$\sum \lvert c(t)\rvert, t > 1.9835$	56.3%

Tab. 2: Feature importance for multi-feature DT

Feature	Importance
$\sum \lvert d(t)\rvert, t \in [0.9252, 1.9835]$	61.3%
$\sum \lvert d(t)\rvert, t > 1.9835$	36.9%
$\sum \lvert b(t)\rvert, t \in [0.9252, 1.9835]$	0.4%
$\sum \lvert b(t)\rvert, t > 1.9835$	0.3%
$\sum \lvert c(t)\rvert, t \in [0.9252, 1.9835]$	0.6%
$\sum \lvert c(t)\rvert, t > 1.9835$	0.5%

As the two features selected in Figure 3 suggests separability, as well as suggesting the highest feature importance according to Tables 1 and 2, a DT classifier was evaluated using only these features with a 10%-90% test-train split. This produced a 98.8% classification accuracy, as compared to the 98.1% accuracy using the Naive Bayes method[3], with confusion matrices seen in Figure 5 (d) and (e) respectively.

Conclusions

We demonstrated the simplicity of the classification on the given dataset: a baseline accuracy of 90.9% was achieved using a manually designed single-feature classifier, where the singular feature was an aggregate which did not require limitation to a narrower time range. In general, aggregates based on the $d(t)$ variable produced the most accurate DT classifiers. Additionally, the data poses a limitation: since the events were recorded individually, it does not providing any suggestion as to how concurrent events might interact on a single cable. It is unclear if concurrent events would stack or if they would cause destructive interference. Future work should examine implementation on more representative datasets, particularly field data[4].

Further improvements upon this method could include limiting depth or pruning the DT. Although the DT classifier produced a higher accuracy than had been previously achieved[3], the anomalous samples inspected, such as the one presented in Figure 4, suggest that this method may not be effective in field applications where events were not so neatly separated, and where different event types may have vastly different lengths of signatures. Thus, future research into localization for events is needed. For this task, CNN[10]–[12] demonstrates promise in similar applications. Such unsupervised algorithms would additionally provide better automation, and may even uncover event types not previously considered.

Acknowledgements

This work was supported by Mitacs under grant IT23512 and Ciena Corp. We wish to thank Todd Morris of Ciena for interesting discussions and Vincent Lemaire (Orange) for his contribution in understanding of the SOP data and methods from his research[3], additionally providing a more exhaustive list of aggregates of interest, as well as the confusion matrix.

References

[1] X. Liu, B. Jin, Q. Bai, Y. Wang, D. Wang, and Y. Wang, "Distributed fiber-optic sensors for vibration detection",

Sensors, vol. 16, p. 1164, Jul. 2016. DOI: 10.3390/s16081164.

[2] J. Teixeira, I. Leite, S. Silva, and O. FrazÃ£o, "Advanced fiber-optic acoustic sensors", *Photonic Sensors*, vol. 4, Sep. 2014. DOI: 10.1007/s13320-014-0148-5.

[3] V. Lemaire, F. Boitier, J. Pesic, A. Bondu, S. Ragot, and F. Clerot, "Proactive fiber break detection based on quaternion time series and automatic variable selection from relational data", in Jan. 2020, pp. 26–42, ISBN: 978-3-030-39097-6. DOI: 10.1007/978-3-030-39098-3_3.

[4] K. Guan, J. Simsarian, F. Boitier, D. Kilper, J. Pesic, and M. Sherman, "Efficient classification of polarization events based on field measurements", Jan. 2020, Th3D.7. DOI: 10.1364/OFC.2020.Th3D.7.

[5] J. PESIC, E. Rouzic, N. Brochier, and L. Dupont, "Proactive restoration of optical links based on the classification of events", Mar. 2011, pp. 1–6.

[6] Z. Zhan, M. Cantono, V. Kamalov, *et al.*, "Optical polarization-based seismic and water wave sensing on transoceanic cables", *Science*, vol. 371, pp. 931–936, Feb. 2021. DOI: 10.1126/science.abe6648.

[7] H. Wu, Z. Wang, F. Peng, *et al.*, "Field test of a fully distributed fiber optic intrusion detection system for long-distance security monitoring of national borderline", *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 9157, Jun. 2014. DOI: 10.1117/12.2058504.

[8] E. Marszalec, J. Marszalec, and R. Romaniuk, "On the application of fibre optics in the development of sensor systems of intelligent robots: A review", *International Journal of Optoelectronics*, vol. 4, pp. 313–325, May 1989.

[9] V. Lemaire, C. Salperwyck, and A. Bondu, "A survey on supervised classification on data streams", *Lecture Notes in Business Information Processing*, Mar. 2015. DOI: 10.1007/978-3-319-17551-5_4.

[10] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Deep learning for time series classification: A review", *Data Mining and Knowledge Discovery*, vol. 33, Jul. 2019. DOI: 10.1007/s10618-019-00619-1.

[11] H. Ren, Z. Hou, B. Vyakaranam, H. Wang, and P. Etingov, "Power system event classification and localization using a convolutional neural network", *Frontiers in Energy Research*, vol. 8, Nov. 2020. DOI: 10.3389/fenrg.2020.607826.

[12] T. Amarbayasgalan, V.-H. Pham, N. Theera-Umpon, and K. Ryu, "Unsupervised anomaly detection approach for time-series in multi-domains using deep reconstruction error", *Symmetry*, vol. 12, p. 1251, Jul. 2020. DOI: 10.3390/sym12081251.