Disturbance pattern recognition based on an ALSTM in a long-distance φ-OTDR sensing system

Xue Chen¹ | Chengjin Xu²

¹Faculty of Medicine, University of Cologne, Cologne, Germany
²Institute of Computer Science, University of Bonn, Bonn, Germany

Abstract
In this article, a new pattern recognition method for disturbance signals detected by phase-sensitive optical time domain reflectometry (φ-OTDR) distributed optical fiber sensing systems is proposed. Currently, most of the disturbance signal recognition methods for φ-OTDR exploit the global features of disturbance signals as the basis of classification, neglect the local details of disturbance signals, and thus have poor performances on long-distance monitoring tasks. In the method proposed in this article, an adaptive denoising method based on spectral subtraction is utilized to enhance signal features. For each frame of disturbance signals, Mel-frequency cepstral coefficients are extracted as frequency-domain features, while short-time energy ratio and short-time level crossing rate are extracted as time-domain features. An attention-based long short-term memory network is exploited as a classifier to recognize different types of disturbance signals. Experiments show the proposed disturbance recognition method can achieve a classification accuracy of 94.3% with five typical disturbances, namely, walking, digging, vehicle-passing, climbing, and heavy rain, at ranges of up to 50 km.

KEYWORDS
attention mechanism, disturbance recognition, long short-term memory, MFCC, φ-OTDR

1 | INTRODUCTION
Distributed optical fiber sensing systems (DOFSs) have been widely applied in many fields such as perimeter security,¹ oil and gas pipeline monitoring,² and so on, due to its advantage of high sensitivity,³ simple structure, long detection distance, and no power supply requirement. Among various DOFSs, the distributed optical fiber disturbance sensing system based on phase-sensitive optical time domain reflectometry (φ-OTDR) can precisely detect and position multiple disturbances at the same time with only one fiber⁴-⁶ and thus become the new hotspot in the present DOFS research field.

In addition to disturbance detection and positioning, much work on pattern recognition of disturbances has been carried out, in order to improve the performances of φ-OTDR systems. An improved classification algorithm based on multiscale wavelet packet Shannon entropy and neural networks for disturbances in φ-OTDR systems achieves an 85% classification accuracy.⁷ A morphological feature extraction method for intrusion event recognition increases the recognition accuracy to over 90% with three types of disturbances in a φ-OTDR system with a length of 10 km.⁸ These methods utilize the global features of disturbance signals as the basis of classification, but ignore the local details of disturbance signals. With the development of deep learning technology, deep learning-based classification methods in φ-OTDR systems are proposed and achieve better performances on long-distance monitoring.⁹,¹⁰ However, in the existing methods, temporal information in disturbance signals is missing since these methods convert disturbance signals detected by φ-OTDR systems into corresponding image data as input features.

In this article, a novel recognition method in which a special attention-based long short-term memory (ALSTM) network is utilized as a classifier is introduced in φ-OTDR sensing systems for the first time. We evaluate and compare our method and several existing disturbance recognition methods on a disturbance signal dataset which was collected by an φ-OTDR distributed optical fiber sensing system with a monitoring distance of more than 50 km. Experiment results show that our method outperforms other methods in the same dataset.
The main contributions of our work can be summarized as follows:

- Referring to advanced speech recognition approaches,11 this method exploits an adaptive spectrum subtraction to denoise disturbance signals and extract Mel-frequency cepstral coefficient (MFCC) features as the input of long short-term memory (LSTM) classifier. The model is able to capture both temporal information and local details of disturbance signals and have been proved to be effective.
- In a segment of disturbance signal, the part which time-domain feature values are relatively high or have obvious changes should be paid more attention to in the recognition model. Thus, we introduce time-domain features into our basic LSTM model by a special attention mechanism to improve model performance.
- Experimental results indicate that our approach can achieve a higher classification accuracy in a same dataset compared with several existing methods.7,8 Meanwhile, we develop a heterogeneous multicore CPU and GPU architecture to implement our approach and achieve online monitoring. We strongly believe that our approach can be generalized in various monitoring systems.

The rest of this article is structured as follows: Section 2 gives an introduction of signal processing methods in our system; Section 3 gives a description of the structure of our proposed classification model; Section 4 introduces the setup of our q-OTDR system which is used to sample disturbance signals and presents extensive experiments to justify the effectiveness of our proposals; and Section 5 summarizes this work.

2 | SIGNAL PROCESSING

2.1 | Spectral subtraction

Spectral subtraction is one of the most effective speech enhancement methods.12 Spectral subtraction has the characteristics of few constraints, specific physical meaning and small computation time, and thus can easily realize fast processing and achieve high signal-noise rates, which makes it widely used.

This approach uses the characteristics of the irrelevance between the noise and the pure signal in the noisy signal. Assuming that the noise is statistical stationary, the noise spectrum of the noisy signal can be estimated by the Fourier spectrum of the voiceless signal. The noise spectrum is subtracted from the noisy speech spectrum so as to obtain the estimated value of the pure speech spectrum. This additive model can be expressed as:

\[ X_w(\omega) = S_w(\omega) + N_w(\omega) \]

where \( X_w(\omega) \) is the short-time Fourier spectrum of the small frame of the noisy signal, \( S_w(\omega) \) and \( N_w(\omega) \) are the short-time Fourier spectra of noise component and effective component of the signal frame, respectively. Since \( S_w(\omega) \) and \( N_w(\omega) \) are independent of each other in the frequency spectrum, the cross-correlation statistical mean is zero. The short-time power spectrum of the pure signal can be estimated as:

\[ |\hat{S}_w(\omega)|^2 = |Y_w(\omega)|^2 - \mathbb{E}[|N_w(\omega)|^2] \]

where \( \hat{S}_w(\omega) \) is the short-time Fourier spectrum estimation of the pure signal. At final, we can use inverse Fourier transform to obtain the denoised waveform with the restored phase information.

\[ \hat{s}(m) = \text{IFFT} \left[ |\hat{S}_w(\omega)| e^{j\phi(\omega)} \right] \]

Since the noise is locally stationary, in our approach, the noise power spectrum of the disturbance signal can be estimated from the mean short-time power spectrum of non-vibrational signal frames which are within the last 0.6 second just before the occurrence of the disturbance in the disturbance position.

2.2 | Feature analysis

In audio classification tasks, one of the key steps is extracting effective discriminative features from the signal. The classification accuracy can be improved if the selected features can express the temporal and spectral characteristics of the signal. At present, short-time energy (STE), zero-crossing rate (ZCR), and MFCCs are the most commonly used time-domain features and frequency domain features respectively, which have been proved to be effective on audio and vibration signal classification tasks. Disturbance signal can also be treated as a special type of audio sequence because its frequency distribution is within the frequency range of human hearing.13 Inspired by some works in audio classification,14 we select these three kinds of signal features as the classification basis in our system.

2.2.1 | Mel-frequency cepstral coefficients

MFCCs are cepstrum parameters extracted from the Mel scale frequency domain.15 As Figure 1 shows, the Mel frequency scale describes the nonlinear characteristics of the
frequency perception of the ear. The transformation relationship between MFCC and linear frequency is calculated as:

$$\text{Mel}(f) = 2595 \times \log(1 + f/700)$$

We can extract MFCCs through the following five steps:

1. Segment signal into frames and make window processing.
2. Compute the frequency spectrum of each signal frame by discrete Fourier transform.
3. Filter the frequency spectrum of the signal through Mel-Scale filter banks.
4. Compute the logarithmic energy of the output of each filter group.
5. Calculate MFCCs through discrete cosine transform which is expressed as:

$$c_j = \sqrt{\frac{2}{M}} \sum_{j=1}^{L} m_j \cos\left(\frac{\pi j}{M} (j - 0.5)\right), i = 1, \ldots, I$$

where $M$ is the number of filter banks, $L$ is the set length of the cepstrum, and $I$ is the number of data points in each frame. Here, we set $M = 24$, $L = 12$, and $I = 128$.

In our implementation, disturbance signals are segmented into signal frames by Hamming windows with a 50% inter-frame overlap. The system sampling rate is 5000 Hz and the time length of each frame is 64 ms. In order to obtain the dynamic information in disturbance signals, we compute the first and second derivatives of MFCCs and combine them into a feature vector.

**2.2.2 | STE and ZCR**

The short-time energy and zero-crossing rate are the basic time-domain features for sequence signal recognition. STE can reflect the transient change in the fluctuation intensity of the signal over time, while ZCR which reflects the frequency information of the signal in some degree can be regarded as a rough estimation of spectral characteristics. STE and ZCR can be used to not only signal recognition but also signal detection, due to their less computation.

In our approach, we calculate the STE and ZCR of each optical signal frame at each position on the sensing fiber in real time as the basis of online disturbance detection. Moreover, for each frame of disturbance signals, we concatenate STE and ZCR into a feature vector as the input of the attention-based LSTM model.

**3 | PRINCIPLE OF CLASSIFICATION**

**3.1 | Long short-term memory**

Recurrent neural network (RNN) is one of the most common types of feed-forward neural networks, which is suitable for processing sequence data. However, basic RNN has the gradient vanishing or exploding problems. In order to overcome this shortcoming, LSTM was proposed and achieved superior performance on audio recognition tasks. As Figure 2 shows, a standard LSTM contains three gates and a cell memory state.

More concretely, each cell in LSTM can be computed as follows:

$$X = \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix}$$

$$f_t = \sigma(W_f x + b_f)$$

$$i_t = \sigma(W_i x + b_i)$$

$$o_t = \sigma(W_o x + b_o)$$

$$c_t = f_t \times c_{t-1} + i_t \times \tanh(W_c x + b_c)$$

$$h_t = o_t \times \tanh(c_t)$$

where $W_f$, $W_i$, $W_o \in \mathbb{R}^{d \times 2d}$ are the weight matrices of the input, forget and output gates respectively, and $b_f$, $b_i$, $b_o \in \mathbb{R}^d$ are the corresponding biases of LSTM ($d$ is the dimension of input features). $\sigma$ is the sigmoid function and $\times$ denotes element-wise multiplication. $x_t$ is the input feature vector of LSTM cell unit, while $h_t$ is the vector of the hidden layer.

In our system, the input vector $x_t$ denotes the MFCC feature of a disturbance signal frame. We regard the last hidden vector $h_N$ as the representation of the type of the disturbance signal and input $h_N$ into a softmax classifier. Figure 3 illustrates our basic classification model based on a standard LSTM.

**3.2 | Attention-based long short-term memory**

In a segment of disturbance signal, the part that contains the key information about signal characteristics often has...
relatively high time-domain features or time-domain features have an obvious change in this part. Thus, we propose to design an attention mechanism that can capture the key part of the disturbance signal. Figure 4 illustrates the architecture of our proposed ALSTM-based classification model.

Let \( H \in \mathbb{R}^{d \times N} \) denote the matrix comprising of hidden vectors \( [h_1, h_2, \cdots, h_N] \) in the LSTM network, and \( N \) is the number of signal frames in a disturbance signal sample. Moreover, \( v_i \) stands for the time-domain features of the \( i \)th frame of the disturbance signal sample, containing STE and ZCR, and \( \alpha \) is an attention weight vector generated by the attention mechanism. The final signal feature representation can be obtained by:

\[
M = \tanh \left( \frac{W_h H}{W_v V} \right) \alpha = \text{softmax}(w^T M) \\
r = Ha^T
\]

where \( M \in \mathbb{R}^{(d+d_1) \times N} \), \( \alpha \in \mathbb{R}^N \), \( W_h \in \mathbb{R}^{d \times d} \), \( W_v \in \mathbb{R}^{d_1 \times d_1} \), \( w \in \mathbb{R}^{d+d_1} \) are projection parameters to be learned during training and \( V = [v_1, v_2, \cdots, v_N] \) is a matrix containing time-domain features of all frames in the disturbance signal sample. \( r \in \mathbb{R}^d \) is the feature representation of the disturbance signal and will be input into a softmax classifier to output conditional probability distribution.

\[
y = \text{softmax}(W_s r + b_s)
\]

where \( W_s \) and \( b_s \) are the parameters for the softmax classifier.

4 | EXPERIMENT SETUP AND DISCUSSION

4.1 | Experiment setup

The architecture of our experimental system is shown in Figure 5. The lasers emitted from a narrow linewidth laser with a line width of 0.1 kHz, a wavelength of 1550 nm, and an output power of 40 mW are modulated into continuous optical pulses with a pulse width of 200 ns and a frequency of 2 kHz by an acoustic-optic modulator and a function generator. The optical pulses are amplified by an erbium-doped fiber amplifier (EDFA) and injected into a sensing fiber with a total length of 50 km through a circulator. The backward Rayleigh scattering light interferes in the sensing fiber. The interference light is detected by a photoelectric detector after being amplified by an EDFA and finally collected by a data acquisition card with a sampling rate of 25 MHz. A monitoring software is developed for real-time signal processing.

When a disturbance occurs on the sensing fiber, the phase change of the transmission light in the disturbance position leads to the phase change of the corresponding backscattered light due to elastic-optical effect. Thus, the monitoring software can detect the disturbance position by analyzing the change of the interference light intensity in each position and furthermore recognize the types of disturbances.

4.2 | Disturbance simulation experiments

Disturbance simulation experiments were carried out at two location of the sensing optical fiber in our system, namely 20 and 50 km. A 100-m-long optical fiber was buried in the underground of 20-cm-depth at the location of 20 km. We thumped the ground with a shovel, paced and drove back and forth at this position to generate three types of signal samples for digging, walking and vehicle-passing. At the location of 50 km, we hang a 100-m-long optical fiber on a fence, and then shook the fence to simulate climbing events. Moreover, we collected the heavy rain signals at the location of 50 km in rainy weather.

For each type of disturbances, we conducted 200 simulation experiments and the monitoring system was responsible for detecting and restoring disturbance signals. During the experiments, we collected the optical signals at five consecutive position nodes around the disturbance position as data samples of each type of disturbance signals. In this way, 5000 data samples were generated in total. The time length and the total sampling point number of a single data sample
were set as 2.4 seconds and 4800, respectively, ensuring that our system can respond to the known disturbances in time.

4.3 | Denoising experiments

To enhance the time-frequency domain characteristics of disturbance signals and improve the classification rates of disturbance events, we apply the method introduced in Section 2.1 to denoise disturbance signals. Figure 6 shows the waveforms and the spectrograms of a typical digging signal before and after being denoised.

It can be seen from Figure 6A,C that spectrum subtraction can significantly improve the signal quality of the digging signal on the premise of not weakening the intensity and time-domain characteristics of the pure signal. Figure 6B,D illustrates that the wide-band background noise in the original signal is effectively suppressed and other noise with multiple harmonics generated during the process of data acquisition is completely removed after noise reduction. The typical waveforms of all five types of denoised disturbance signals are shown in Figure 7.

4.4 | Classification experiments

We apply the feature extraction methods introduced in Section 2.2 and the classification models proposed in Sections 3.1 and 3.2 to disturbance classification in φ-OTDR distributed optical fiber sensing systems.
Among 1000 samples of each disturbance mode, 750 samples generated in the random selected 150 trials are chosen as a part of the training set and 250 samples generated in other 50 trials are taken as testing data. In our experiments, we first extracted MFCCs, STE, and ZTR for each frame of each sample. The feature extraction process was encoded in Python and took less than 0.1 second for each sample. And then we trained our proposed model for 30 epoch with a batch size of 75, an initial learning rate as 0.001, a dropout value as 0.5 and a momentum factor of 0.9. The training processes were performed in Tensorflow, running on an NVIDIA GTX TITAN X Pascal GPU with 12-GB onboard memory. Each network was tested on the testing phase every 25 iterations.

**FIGURE 7** The typical waveforms of all five types of denoised disturbance signals. A, Walking; B, digging; C, vehicle-passing; D, climbing; and E, heavy rain
Figure 8 shows the train loss curves and test accuracy curves of LSTM model and ALSTM model trained with a same dataset.

It can be seen from Figure 8 that compared with the LSTM-based classification model, the ALSTM-based classification model has a faster converges speed, a lower training loss and a higher classification rate. Through 35 epochs, the test accuracy curve and train loss curve of the LSTM model gradually converged to 90.6% and 0.35. Meanwhile, it only took 20 epochs for the test accuracy curve and the train loss curve of the ALSTM-based classification model to converge to 94.3% and 0.28, respectively. This was because the attention mechanism made the model to focus on the key parts of disturbance signals faster, which expressed the main signal characteristics. The concrete classification results of the LSTM-based classification model and ALSTM-based classification model on the test set are respectively shown in Tables 1 and 2.

In Tables 1 and 2, each row represents the classification outcomes of the LSTM model or the ALSTM model for testing samples of the corresponding type of disturbance signals. These outcomes show both the LSTM model and the ALSTM model can effectively distinguish between instantaneous behaviors (walking and digging) and long-time behaviors (vehicle passing, climbing and heavy rain), and achieve high classification rates of over 90%. However, the LSTM model might be confused by similar disturbance modes as shown in Table 1. The experiment results in Table 2 indicate that this problem was significantly alleviated through the attention mechanism.

We compare our proposed methods with several existing methods in the same dataset of disturbance signals, including an intrusion event recognition method based on morphologic feature extraction (MFE) and a vibration event recognition method based on convolutional neural networks. The recognition accuracies and recognition time for a single data sample of different recognition methods are shown in Table 3.

Table 3 illustrates that the ALSTM-based recognition method has the best performance, but its computing time for a single sample is longer than other methods. In order to achieve online disturbance detection and recognition in the φ-OTDR distributed optical fiber sensing system, we develop a heterogeneous multicore CPU and GPU architecture to implement our proposed method. In this architecture, CPU executes global logic control and disturbance detection task, while GPU is...
TABLE 3  Recognition accuracies and recognition time of different recognition methods

| Models  | Recognition accuracy (%) | Recognition time (%) |
|---------|--------------------------|----------------------|
| MFE     | 88.1                     | 0.21                 |
| CNN     | 89.9                     | 0.49                 |
| LSTM    | 90.6                     | 0.87                 |
| ALSTM   | 94.3                     | 0.91                 |

Abbreviations: ALSTM, attention-based long short-term memory; CNN, convolutional neural networks; LSTM, long short-term memory; MFE, morphologic feature extraction.

responsible for identifying the detected disturbance signals. Due to no need of sharing parameters, these two processes can be carried out independently at the same time. In this way, we can ensure that a slower recognition speed will not lead to data overflowing during the acquisition process in our system, or have other negative influences on online monitoring of our system.

5 | CONCLUSION

In this article, we have proposed a novel recognition method based on an attention LSTM for disturbance events in q-OTDR distributed optical fiber sensing systems. Disturbance signals are denoised by spectral subtraction to enhance the characteristics of signals. STE, ZCR, and MFCCs extracted from each frame of disturbance signals are regarded as the classification basis to be input into an attention-based LSTM model.

The core idea for our method is to utilize the LSTM model to capture both temporal information and local details of disturbance signals and exploit the attention mechanism to lead our model to find out the key parts of disturbance signals which fully express the signal characteristics. Experimental results show that our proposed ALSTM based recognition method achieves a high classification accuracy of 94.3% with five typical disturbances, namely, walking, digging, vehicle-passing, climbing, and heavy rain, at ranges of up to 50 km, better than the performance of the LSTM based method and other two advanced methods on the same dataset. In order to ensure the reliability of our online monitoring system, we develop a heterogeneous multicores CPU and GPU architecture to implement our methods to process signals in real time. Moreover, we strongly believe that our proposed methods can be generalized in other types of optical fiber sensing systems and various online monitoring systems for some one-dimensional time series signals, such as audio signals and mechanical vibration signals.

ORCID

Chengjin Xu https://orcid.org/0000-0002-4942-9049

REFERENCES

[1] Wu H, Wang Z, Peng F, et al; Field test of a fully distributed fiber optic intrusion detection system for long-distance security monitoring of national borderline. 23rd International Conference on Optical Fibre Sensors. 2014;9157:915790.

[2] Tejedor J, Martins HF, Piote D. Toward prevention of pipeline integrity threats using a smart fiber-optic surveillance system. J Lightwave Technol. 2016;34(19):4445-4453.

[3] Bao X, Chen L. Recent progress in distributed fiber optic sensors. Sensors. 2012;12(7):8601-8639.

[4] Taylor HF, Juarez JC. Field test of a distributed fiber-optic intrusion sensor system for long perimeters. Appl Optics. 2007;46(11):1968.

[5] Masoudi A, Belal M, Newson TP. A distributed optical fibre dynamic strain sensor based on phase-OTDR. Meas Sci Technol. 2013;24(8):085204.

[6] Xu C, Guan J, Bao M. Pattern recognition based on enhanced multi-feature parameters for vibration events in q-OTDR distributed optical fiber sensing system. Microw Opt Technol Lett. 2017;59(12):3134-3141.

[7] Wang B, Pi S, Sun Q. Improved wavelet packet classification algorithm for vibrational intrusions in distributed fiber-optic monitoring systems. Opt Eng. 2015;54(5):055104.

[8] Sun Q, Feng H, Yan X. Recognition of a phase-sensitivity OTDR sensing system based on morphologic feature extraction. Sensors. 2015;15(7):15179-15197.

[9] Xu C, Guan J, Bao M. Pattern recognition based on time-frequency analysis and convolutional neural networks for vibrational events in q-OTDR. Opt Eng. 2018;57(1):1.

[10] Aktas M, Akgun T, Demircin MU. Deep learning based threat classification in distributed acoustic sensing systems, In: Signal Processing and Communications Applications Conference IEEE; 2017.

[11] Hori T, Chen Z, Erdogan H. Multi-microphone speech recognition integrating beamforming, robust feature extraction, and advanced DNN/RNN backend. Comput Speech Lang. 2017;46:401-418.

[12] Palival K, Jcicki K, Schwerin B. Single-channel speech enhancement using spectral subtraction in the short-time modulation domain. Speech Commun. 2010;52(5):450-475.

[13] Tsau E. Advanced Features and Feature Selection Methods for Vibration and Audio Signal Classification [Ph.D. dissertation]. California: Dept. Elect. Eng., Southern California Univ.; 2012.

[14] Dahake PP, Shaw K, Malathi P, Speaker dependent speech emotion recognition using MFCC and support vector machine, In: International Conference on Automatic Control and Dynamic Optimization Techniques IEEE; 2017:1080-1084.

[15] Carey MJ, Parris ES, Thomas HL. A comparison of features for speech, music discrimination, In: IEEE International Conference on Acoustics, Speech, and Signal Processing; 1999:149-152.

[16] Saunders J. Real-time discrimination of broadcast speech/music, In: IEEE International Conference on Acoustics, Speech, and Signal Processing; 1996:993-996.

[17] Greff K, Srivastava RK, Koutnik J. LSTM: a search space odyssey. IEEE Trans Neural Netw Learn Syst. 2016;28(10):2222-2232.

How to cite this article: Chen X, Xu C. Disturbance pattern recognition based on an ALSTM in a long-distance q-OTDR sensing system. Microw Opt Technol Lett. 2020;62:168–175. https://doi.org/10.1002/mop.32025