

A Harmful-Intrusion Detection Method Based on Background Reconstruction and Two-Dimensional K-S Test in an Optical Fiber Pre-Warning System

Fukun BI¹, Tong ZHENG^{1*}, Hongquan QU¹, and Liping PANG²

¹College of Information Engineering, North China University of Technology, Beijing, 100144, China

²School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing, 100191, China

*Corresponding author: Tong ZHENG E-mail: dolly115dolly@163.com

Abstract: The key technology and main difficulty for optical fiber intrusion pre-warning systems (OFIPS) is the extraction of harmful-intrusion signals. After being processed by a phase-sensitive optical time-domain reflectometer (Φ -OTDR), vibration signals can be preliminarily extracted. Generally, these include noises and intrusions. Here, intrusions can be divided into harmful and harmless intrusions. With respect to the close study of signal characteristics, an effective extraction method of harmful intrusion is proposed in the paper. Firstly, in the part of the background reconstruction, all intrusion signals are first detected by a constant false alarm rate (CFAR). We then reconstruct the backgrounds by extracting two-part information of alarm points, time and amplitude. This ensures that the detection background consists of intrusion signals. Secondly, in the part of the two-dimensional Kolmogorov-Smirnov (K-S) test, in order to extract harmful ones from all extracted intrusions, we design a separation method. It is based on the signal characteristics of harmful intrusion, which are shorter time interval and higher amplitude. In the actual OFIPS, the detection method is used in some typical scenes, which includes a lot of harmless intrusions, for example construction sites and busy roads. Results show that we can effectively extract harmful intrusions.

Keywords: Optical fiber intrusion pre-warning; extraction of harmful-intrusion signals; two-dimensional K-S test

Citation: Fukun BI, Tong ZHENG, Hongquan QU, and Liping PANG, "A Harmful-Intrusion Detection Method Based on Background Reconstruction and Two-Dimensional K-S Test in an Optical Fiber Pre-Warning System," *Photonic Sensors*, 2016, 6(2): 143–152.

1. Introduction

Intrusion prediction for pipelines (IPP) is used to monitor the transport process of petroleum and natural gas. It is well known that leakage pre-warning can prevent many fatal losses. Optical fibers buried along the pipeline are generally used to transmit data. Moreover, the vibration measurement of these optical fibers is an effective and convenient

way to detect intrusion. In other words, we can locate harmful intrusions and predict damage with the proposed method [1–5].

At present, there are two types of regime in the field of optic fiber intrusion, namely reflection and interference. The Mach-Zehnder (MZ) method is an interference method and applied to the detection of synthesis signals across a whole optical fiber. However, the detection performance will be rapidly

Received: 15 December 2015 / Revised: 2 March 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

DOI: 10.1007/s13320-016-0308-x

Article type: Regular

distorted if there are concurrent vibration signals. Unfortunately, the optical fiber intrusion pre-warning systems (OFIPS) have many concurrent intrusion signals. On the other hand, the phase-sensitive optical time-domain reflectometer (Φ -OTDR), which is a reflection method, can be used to detect concurrent intrusions, even at very small resolutions [6]. However, the Φ -OTDR system also has a limitation in its sensitivity [7–12], thus interference and harmful intrusions are difficult to identify. After a close study, we find that the signal can be divided into three parts: noise, and harmless and harmful intrusions. Here, noise and harmless intrusion are interference. Hence, we need to hierarchically extract different kinds of signals and locate harmful intrusion. The support vector method (SVM) is used to identify different interference in Tian Jin University of China [10].

The background distribution of OFIPS can be divided into two orthogonal Gaussian noises, and the envelope of the synthesis signal obeys a Rayleigh distribution. This character of background distribution is similar to radar signals. For this reason, the constant false alarm rate (CFAR) method in radar detection is introduced here to detect the intrusion signal of optical fibers. In the spatial domain detection, CFAR is divided into several categories, including a mean level CFAR (ML-CFAR), ordered statistics CFAR (OS-CFAR), and adaptive CFAR (A-CFAR). On the one hand, optimal statistics of OS-CFAR depend on prior information [13]. In actual conditions, the prior information is difficult to obtain. On the other hand, A-CFAR can adapt to a variable background, but the running time for this is quite long due to a complicated algorithm [14]. However, CA-CFAR, categorized within ML-CFAR, has a simple algorithm and outstanding performance in a homogeneous background. In view of these facts, the CA-CFAR is used to detect vibration signals [15–17]. In our system, all intrusion signals, including harmless and harmful intrusions, can be

detected by the CFAR method. It is obvious that the extraction of harmful intrusion is a significant part of this paper.

The proposed method includes two parts, namely background reconstruction and a two-dimensional K-S test. Harmful intrusion signals can be properly extracted. As harmful intrusion exists in all intrusion signals, we plan to reconstruct the background. The disposal of the background reconstruct ensures that later detected backgrounds consist entirely of intrusion signals. Moreover, the harmful intrusion signals are continuous within a period of time, so we need to detect the system in time dimensions. It is well known that one CFAR detected method in time dimensions is a clutter map. However, this has the shortcoming of self-shielding [17–18] and cannot adapt to a variable background. According to studies, the background of OFIPS is unstable. In view of this, in the proposed method, we firstly employ the Kolmogorov-Smirnov (K-S) test, which belongs to a non-parametric test, to extract harmful-intrusion signals. The statistics of K-S quantify the distance between the empirical distribution function (EDF) of the sample and the cumulative distribution function (CDF) of the reference distribution, or between the EDFs of two samples. Under the K-S test method, statistical information, such as mean, variance, and distribution type, do not need to be calculated before the test. Hence, it has wide applications [19–23]. Compared with the harmless intrusion, a harmful one has a shorter time interval and higher amplitude. According to the character of the harmful intrusion, we design a detected method, a K-S test in two dimensions. The intrusion type, whether harmful or harmless, can be confirmed by this method.

The remainder of the paper is organized as follows. Section 2 gives an introduction of the proposed method. Section 3 is devoted to the background construction and reconstruction. The method of the K-S test in two dimensions is proposed in Section 4. In Section 5, experiments on real data demonstrate the effectiveness of the

proposed method. The discussion and conclusion are provided in Section 6.

2. Summary of the overall process

The sensing signal includes noise, and harmless and harmful intrusions. Here, noise is caused by the self-vibration of the ground, and it is inevitable. Harmless intrusion, which is generated by the construction on the site or vehicles, has no ability to damage the underground pipeline or optical fiber. Harmful intrusion becomes the target of extraction, such as excavation at the ground. The overall detection process of harmful intrusion is shown in Fig. 1.

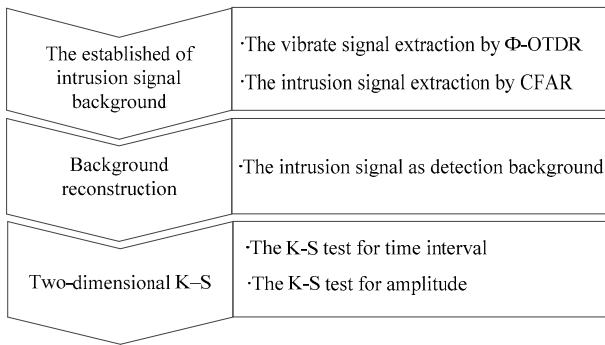


Fig. 1 Overall process of the system.

In the part of background establishment, the sensing signals are preliminarily extracted by Φ -OTDR. It consists of noise, and harmless and harmful intrusions. An intrusion background is established. In addition, all intrusion signals can be obtained, and all interference is eliminated by the CA-CFAR detector.

In view of this, the key point of this paper is the extraction of the harmful intrusion. Firstly, the detection background needs to be reconstructed. In other words, all intrusion signals become new background for later detection. In actual conditions, harmful-intrusion signals have a shorter time interval and higher amplitude than harmless ones. In the view of these characteristics, we apply the K-S test method in two dimensions to detect harmful-intrusion signals.

We provide a two-dimensional K-S test method

to detect the system. K-S tests for the time interval and amplitude are used to extract the harmful intrusion. We do not recognize the intrusion signal as harmful until outcomes of two-dimensional tests both have satisfied corresponding requirements.

3. Construction and reconstruction of an intrusion signal background

3.1 Extraction of vibration signal based on Φ -OTDR

The principal of OFIPS is shown in Fig. 2. If there are any above-ground intrusions, optical fiber buried along the pipeline will perceive the vibration. The vibration can be collected and treated by the hardware system. Finally, the result of intrusion detection needs to be displayed in software. In view of the elastic-optical effect, the refractive index of the optical fiber changes suddenly. This may cause variations in the intensity of back-scattered light. Differences in the intensity of background light measured at different time can be used to locate the vibration [6]. Hence, the Rayleigh backscattered light power is calculated as follows:

$$P_{RB} = 2 \cdot P_i \cdot RC \{1 + \cos[\Phi + \Delta\Phi(t)]\} \quad (1)$$

where P_{RB} is the power of the backscattered light, P_i is the power of source light transmitted from the first end, RC is the backscattered Rayleigh coefficient, Φ is the phase between the rising and trailing edges of light pulse, and $\Delta\Phi(t)$ is the change in phase following time. According to (1), the light power of optical fibers on the intrusion position can change after an aboveground intrusion.

3.2 Extraction of an intrusion signal based on CA-CFAR

The CFAR methods consist of mean-level CFAR, mean level CFAR [14], ordered statistics CFAR [15], adaptive CFAR [16], and so on. Considering the appliance difficulty, the CA-CFAR which belongs to mean-level CFAR is used in our system. The principle of CA-CFAR is shown in Fig. 3. The

estimated background noise level is the mean value of the reference cells. The adaptive threshold is $S=TZ$, where T is the threshold multiplier, and Z represents the noise level. In the detection, the test cell is compared to an estimated threshold. The

sample is judged as an intrusion when it is greater than the threshold. However, the sample is a noise that does not need to be detected. Clearly, the extraction of the intrusion signal by CA-CFAR is effective and efficient.

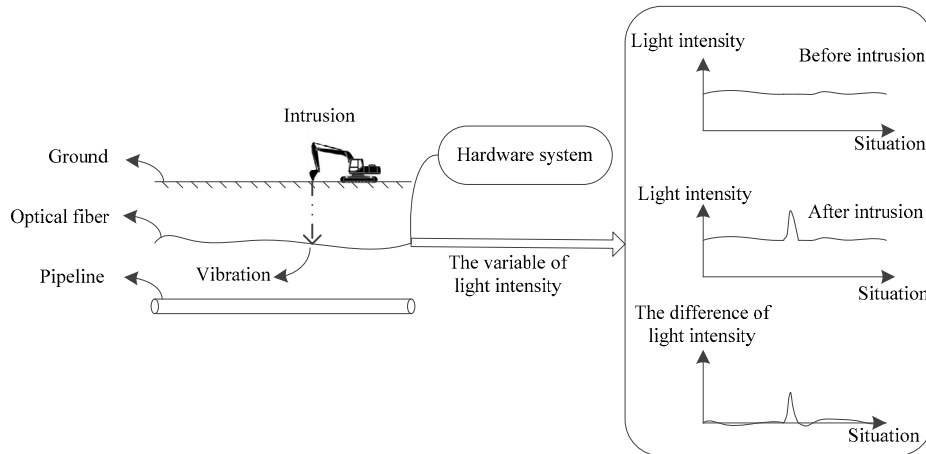


Fig. 2 Principle of OFIPS.

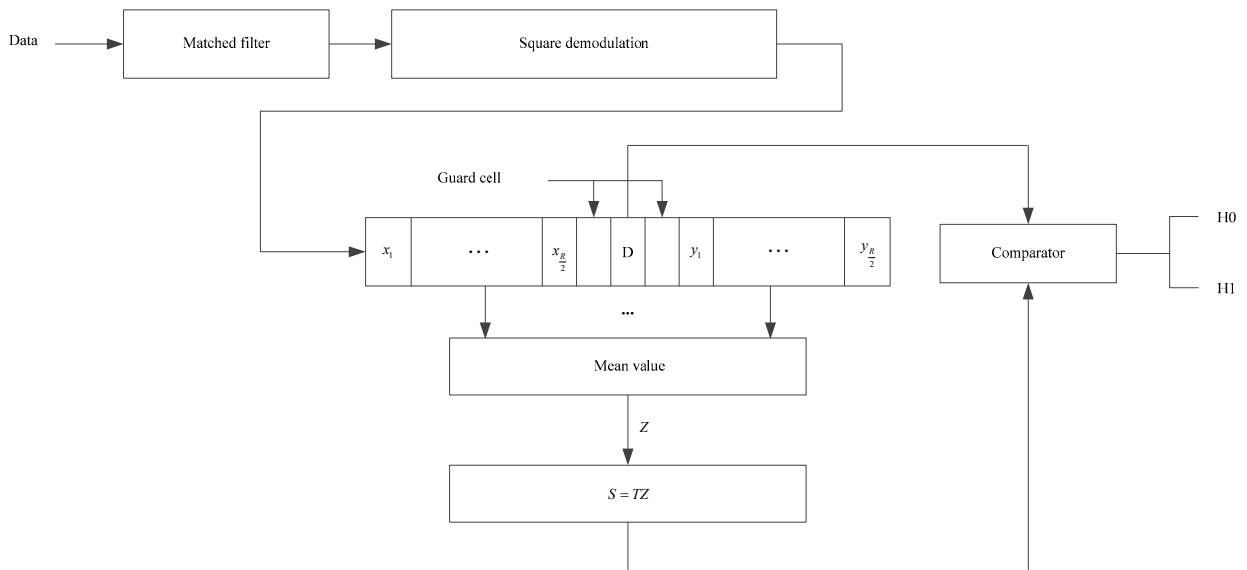


Fig. 3 Principle of CA-CFAR.

3.3 Reconstruction of detection background

The sensing signals can be obtained by Φ -OTDR, and the detection background is established in this way. After the CA-CFAR detection, all intrusion signals are extracted. Intrusions consist of harmless- and harmful-intrusion signals. We need to reconstruct

the background and guarantee that the new one consists of all intrusion signals. The reconstruction method is shown in Fig. 4.

The plane is the detection background before reconstruction in the left side of Fig. 4. The abscissa represents the situation, and the ordinate is time. Intrusions detected by CA-CFAR are marked as “o”. If we assume that every sample set includes four

samples, the method of background reconstruction will collect every four samples extracted by the same situation point and adjacent time point as a sample set. The above process is shown in the right side of Fig.4. The sample numbers in each sample set, N , can be adjusted according to the concrete condition. Finally, we can obtain many sample sets, which consist of intrusion samples. For the latter detection, the information of time interval and amplitude needs to be recorded.

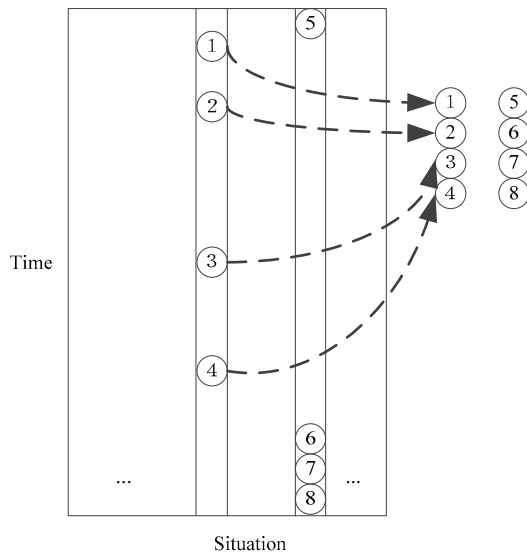


Fig. 4 Schematic diagram of background reconstruction.

4. Harmful-intrusion signal detection based on the K-S test in two dimensions

4.1 Principle of the K-S test

In statistics, the K-S test is a non-parametric test of the equality of distributions. It can be used to compare a sample with a reference probability distribution (a one-sample K-S test) or to compare two samples (a two-sample K-S test). The K-S statistic quantifies the distance between the energy distribution function (EDF) of sample and the cumulative density function (CDF) of the reference distribution, or between the EDFs of two samples. When applying a two-sample K-S test, the distance D between two EDFs $F_{N_1}(x)$ and $R_{N_2}(x)$ is calculated as [19–23]

$$D = \sup |F_{N_1}(x) - R_{N_2}(x)| \quad (2)$$

where N_1 and N_2 are the numbers of data points in the first and second samples, respectively. Clearly, the distance approaches zero following the approach of EDFs. The distance D is used to judge whether the two distributions are homologous.

4.2 Method of the K-S test in two dimensions

Compared with the harmless intrusion, harmful-intrusion signals have the shorter time interval and higher amplitude. According to these characteristics, we design a two-dimensional K-S test method. Firstly, the time-interval K-S test needs to be implemented. If the outcome of the K-S test is that two distributions are homologous, the test sample set is a harmless one. Otherwise, it may be a harmful one, and the amplitude K-S test needs to be executed. Only under the condition that two amplitude distributions are different, the test sample can be recognized as a harmful-intrusion set.

4.2.1 Harmful-intrusion K-S test based on the characteristic of time interval

Considering the characteristics of the harmful intrusion within a shorter time interval, a time-interval detection method has been designed in this paper. The overall process is shown in Fig. 5. Firstly, we perform time-interval K-S tests between every sample set and its adjacent reference harmless set. Secondly, we can judge whether every sample set is harmless or suspected harmful. Finally, the K-S test

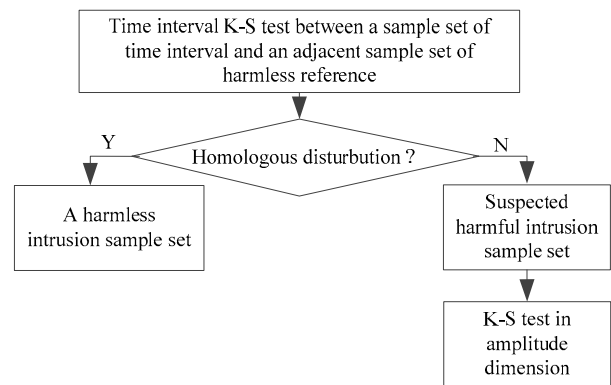


Fig. 5 K-S test in the time interval.

in the amplitude dimensional needs to be employed for every suspected harmful-intrusion sample set.

4.2.2 Harmful-intrusion K-S test based on the amplitude characteristic

In view of the characteristics of the harmful intrusion with higher amplitude, the amplitude K-S test has been designed in the paper. The overall process is shown in Fig. 6. Firstly, we perform amplitude K-S tests between every sample set and its adjacent reference harmless set. Secondly, we can judge whether every sample set is harmless or harmful. If the test outcome is that of a homologous distribution, the testing set is a harmless one. Otherwise, it is a harmful-intrusion sample set. Harmful intrusion can be extracted by this method.

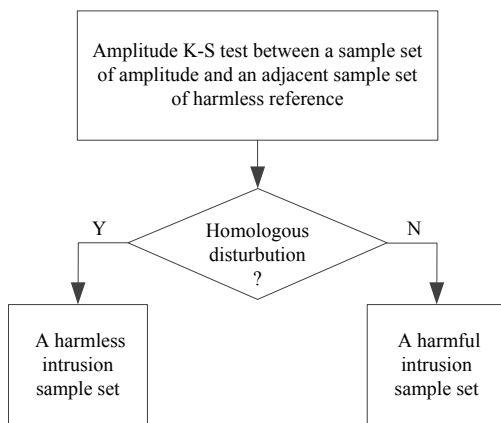


Fig. 6 K-S test in amplitude.

4.2.3 Fusion of the K-S test in time interval and amplitude

According to above two descriptions, the harmful intrusion has the shorter time interval and higher amplitude than harmless one. In other words, if the two outcomes of the K-S test both display with different distributions, the sample set can be recognized as a harmful one. If the sample set is one with a short time interval and low amplitude, the damage to the underground pipeline is negligible. Hence, the condition is defined as a harmless one. On the other hand, a sample set with the long time interval and high amplitude is harmless too. Because

the intrusion is not continuous, damage is generally deemed ignorable. Hence, in the condition that two requirements are satisfied, the testing sample set can be recognized as a harmful-intrusion sample set.

5. Experiments and analysis

The real data for OFIPS were collected from the construction site and busy road of the Da Gang oilfield. The vibration data for roads were obtained on Jin Qi road, and the construction site data were acquired in Yi He village, on the west side of Jin Qi road. The collection hardware system is shown in Fig. 7. This includes pulse light launch, photoelectric conversion, and intrusion-signal acquisition. The actual conditions of fiber burial and tests in the field are shown in Figs. 8 and 9, respectively.

5.1 Confirmation of key parameters

It is clear that the main algorithms of OFIPS are background reconstruction and the two-dimensional K-S test. Here, we confirm the key parameters by simulation and a real data test.

In the part of background reconstruction, we choose 40 as the sample number for each sample set. In other words, the two-dimensional K-S test cannot be implemented until the requirement that each sample set has 40 intrusion points has been first guaranteed. There are two reasons to confirm the parameter. On the one hand, if there are few intrusion points in every sample set, the characteristic of distribution has not been accurately reflected. On the other hand, the time interval of the adjacent vibration point is 1 ms in our system. The running time will be long, and difficulties in intrusion display will emerge, if the intrusion number of each sample set becomes excessive. After exhaustive experimentation, we set the parameter as 40. This parameter can effectively balance two potential threats. Complete distribution can be revealed, and running time is acceptable with this condition.

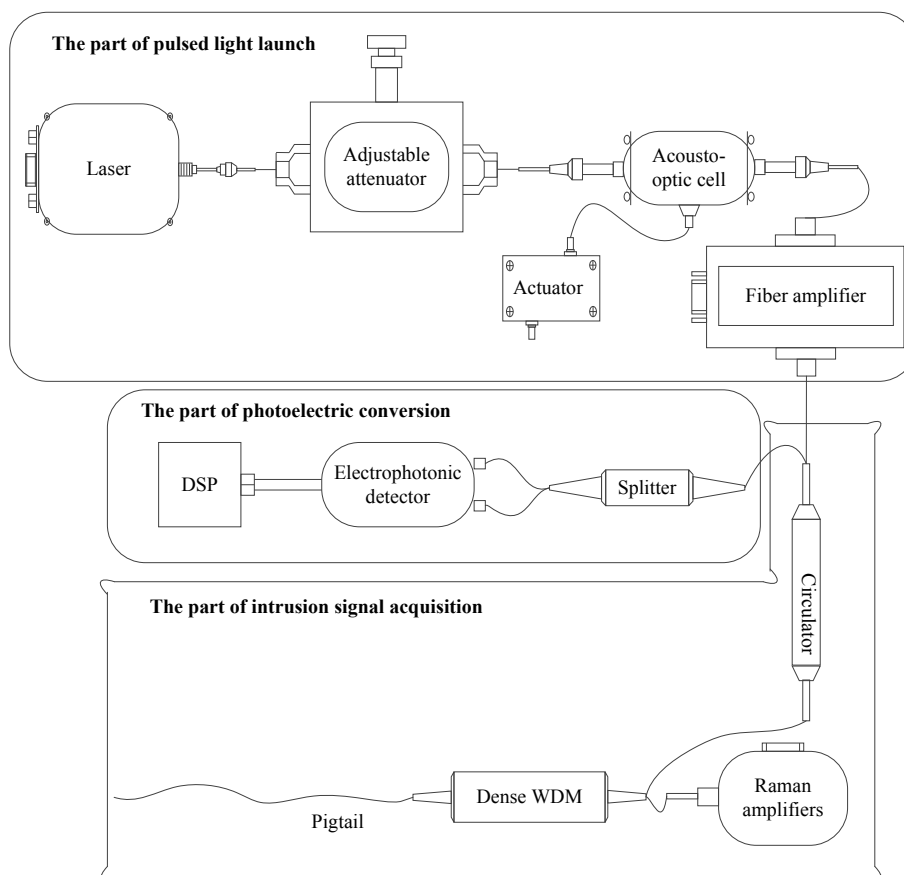


Fig. 7 Hardware system.

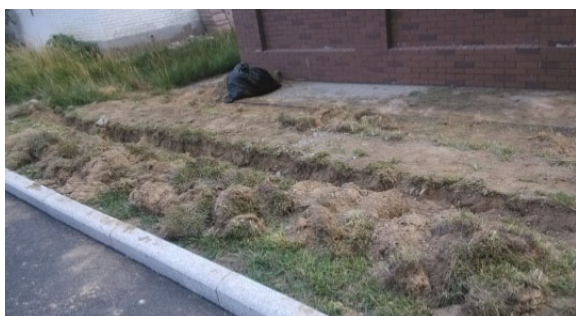


Fig. 8 Fiber buried in the field.



Fig. 9 Field tests.

5.2 Detection test on real data

The testing personnel dug at the construction site using a jackhammer. The test was implemented over a position of 200m in Yi He village. The length of the excavation time was 10 s. After preliminary detection by CFAR, the outcome is shown in Fig. 10(a). Then we extract the harmful intrusion by background reconstruction and K-S test in two dimensions. The final outcome is displayed in Fig. 10(b).

It is clear that there are many intrusion signals after the CA-CFAR detection. The revealed outcome consists of harmless- and harmful-intrusion signals. According to Fig. 10(b), the outcome treated by the proposed method just includes harmful ones. We can intuitively obtain situations and time of harmful intrusions.

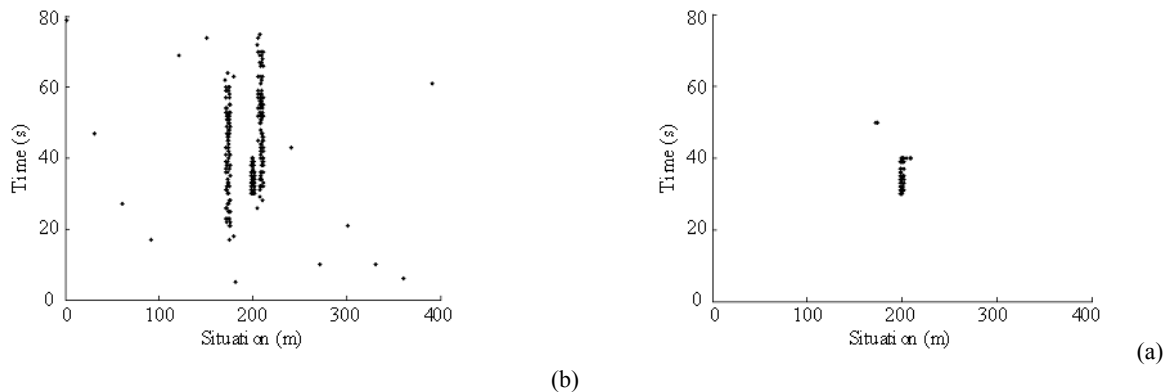


Fig. 10 Detection outcome at the construction site: (a) outcome after the CA-CFAR detecting and (b) outcome after treatment with the proposed method.

On Jin Qi road, testing personnel also dug using a jackhammer. The concrete position was 150m, and the excavation lasted for 10 s. The intrusion outcomes after the CFAR detection and the proposed method are displayed in Figs. 11(a) and 11(b), respectively.

It is clear that there are many diagonal intrusions

after the CA-CFAR detection, caused by driving. Diagonal intrusions are harmless intrusions. Hence, we need to remove harmless intrusions. In general, the outcome treated by the proposed method includes only harmful ones. We can directly obtain the situations and time of harmful intrusions.

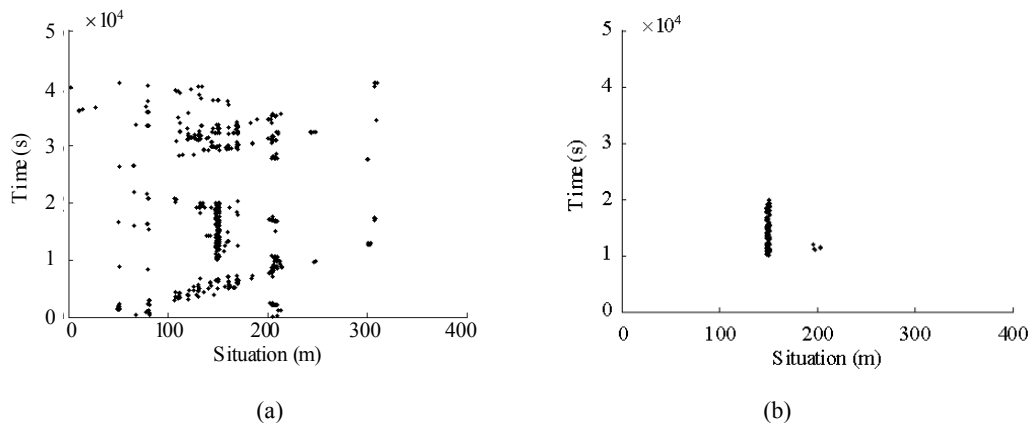


Fig. 11 Detection outcome on the road: (a) outcome after the CA-CFAR detection and (b) outcome after treatment by the proposed method.

6. Discussion and conclusions

Because signals of OFIPS include noise, harmless and harmful intrusions, the method of extracting harmful intrusions from all signals that is provided in the paper is important. We preliminarily extract vibration signals by Φ -OTDR, and the detection background can be established. The CA-CFAR detector is used to extract intrusion signals from all vibration data. The alarm signals consist of both harmless and harmful intrusions.

Hence, the background reconstruction is proposed. Finally, we use the method of the K-S test in two dimensions to detect the harmful-intrusion sample set.

Here, differences in the intensity of background light measured at different time can be used to locate the vibration in Φ -OTDR even at very small resolutions. We apply the CA-CFAR to detect intrusion signals. All alarm points, which include both harmless and harmful intrusions, are detected.

The background reconstruction classifies every adjacent one of 40 intrusion points as a sample set. The time interval and amplitude of each intrusion set should be recorded, and they are used in the later K-S test.

The two-dimensional K-S test includes time interval and amplitude. Firstly, we perform time-interval K-S tests between every sample set and its adjacent set of harmless reference. If the distribution is different, the amplitude K-S test needs to be used. Otherwise, the testing set is a harmless sample set. The amplitude K-S test is similar with a time interval. If amplitude distribution is homologous, the testing set is a harmless one. If not, a harmful-intrusion sample set is extracted.

In sum, we can effectively detect harmful intrusions by the proposed method. The method can be used in many scenes, some of which possess massive harmless intrusions, such as construction sites and busy roads. According to experiments, we find the detection effects to be high quality.

Acknowledgment

The authors are grateful to the anonymous reviewers for their critical and constructive review of the manuscript. This work was funded by the Scientific Pre-research Foundation, the National Natural Science Foundation of China (No. 61571014), and the Scientific Pre-research Foundation of NCUT.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] Z. Qu, H. Feng, Z. Zeng, J. Zhuge, and S. Jin, "A SVM-based pipeline leakage detection and pre-warning system," *Measurement*, 2010, 43(4): 513–519.
- [2] J. Kang and Z. Zou, "Time prediction model for pipeline leakage based on grey relational analysis," *Physics Procedia*, 1989, 25(2): 2019–2024.
- [3] W. Liang, L. Lu, and L. Zhang, "Coupling relations and early-warning for 'equipment chain' in long-distance pipeling," *Mechanical Systems and Signal Processing*, 2013, 41(1–2): 335–347.
- [4] W. Liang, L. Zhang, Q. Xu, and C. Yan, "Gas pipeline leakage detection based on acoustic technology," *Engineering Failure Analysis*, 2013, 31(6): 1–7.
- [5] T. Zhang, Y. Tan, H. Yang, J. Zhao, and X. Zhang, "Locating gas pipeline leakage based on stimulus-response method," *Energy Procedia*, 2014, 61: 207–210.
- [6] Q. Lv, L. Li, H. Wang, Q. Li, and X. Zhong, "Influences of laser on fiber-optical distributed disturbance sensor based on Φ -OTDR," *Infrared and Laser Engineering*, 2014, 43(12): 3918–3923.
- [7] H. F. Martins, S. Martin-Lopez, P. Corredera, M. L. Filograno, O. Frazão, and M. González, "Coherent noise reduction in high visibility phase-sensitive optical time domain reflectometer for distributed sensing of ultrasonic waves," *Journal of Lightwave Technology*, 2013, 31(23): 3631–3637.
- [8] Q. Li, C. Zhang, L. Li, and X. Zhong, "Localization mechanisms and location methods of the disturbance sensor based on phase-sensitive OTDR," *Optik – International Journal for Light and Electron Optics*, 2014, 125(9): 2099–2103.
- [9] Q. Lin, C. Zhang, and C. Li, "Fiber-optic distributed sensor based on phase-sensitive OTDR and wavelet packet transform for multiple disturbances location," *Optik – International Journal for Light and Electron Optics*, 2014, 125(24): 7235–7238.
- [10] Z. Qu, H. Feng, Z. Zeng, J. Zhuge, and S. Jin, "A SVM-based pipeline leakage detection and pre-warning system," *Measurement*, 2010, 43(4): 513–519.
- [11] A. R. Bahrampour and F. Maaoumi, "Resolution enhancement in long pulse OTDR for application in structural health monitoring," *Optical Fiber Technology*, 2010, 16(4): 240–249.
- [12] L. Lu, Y. Song, X. Zhang, and F. Zhu, "Frequency division multiplexing OTDR with fast signal processing," *Optics & Laser Technology*, 2012, 44(7): 2206–2209.
- [13] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE Transactions on Aerospace and Electronic Systems*, 1983, 19(4): 608–621.
- [14] S. D. Himonas and M. Barkat, "Automatic censored CFAR detection for nonhomogeneous environments," *IEEE Transactions on Aerospace*

- and Electronic Systems*, 1992, 28(1): 286–304.
- [15] R. Zhang, W. Sheng, and X. Ma, “Improved switching CFAR detector for non-homogeneous environments,” *Signal Processing*, 2013, 93(1): 35–48.
- [16] G. V. Weinberg, “Management of interference in Pareto CFAR processes using adaptive test cell analysis,” *Signal Processing*, 2014, 104(104): 264–273.
- [17] B. Shi, C. Hao, C. Hou, X. Ma, and C. Peng, “Parametric Rao test for multichannel adaptive detection of range-spread target in partially homogeneous environments,” *Signal Processing*, 2015, 108: 421–429.
- [18] R. Nitzberg, “Clutter map CFAR analysis,” *IEEE Transactions on Aerospace and Electronic Systems*, 1986, 22(4): 419–421.
- [19] X. Wang and V. Makis, “Autoregressive model-based gear shaft fault diagnosis using the Kolmogorov-Smirnov test,” *Journal of Sound and Vibration*, 2009, 327(3–5): 413–423.
- [20] B. Swiderski, S. Osowski, M. Kruk, and J. Kurek, “Texture characterization based on the Kolmogorov-Smirnov distance,” *Expert Systems With Applications*, 2015, 42(1): 503–509.
- [21] J. Rajan, A. J. den Dekker, and J. Sijbers, “A new non-local maximum likelihood estimation method for Rician noise reduction in magnetic resonance images using the Kolmogorov-Smirnov test,” *Signal Processing*, 2014, 103(10): 16–23.
- [22] Z. Drezner and O. Turel, “Normalizing variables with too-frequent values using a Kolmogorov-Smirnov test: a practical approach,” *Computer & Industrial Engineering*, 2011, 61(4): 1240–1244.
- [23] R. Gong and S. Huang, “A Kolmogorov-Smirnov statistic based segmentation approach to learning from imbalanced datasets: with application in property refinance prediction,” *Expert Systems with Applications*, 2012, 39(6): 6192–6200.