WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Practicality of Gas Pipeline Inspection System

Swati Umbre [a], Sujata Gaikwad [b]

[a] *Student, Dr. Babasaheb Ambedkar Technological University, Department of Computer Science and Engineering, College of Engineering Osmanabad, Osmanabad 413501, Maharashtra, India*

[b] *Head of Department, Dr. Babasaheb Ambedkar Technological University, Department of Computer Science and Engineering, College of Engineering Osmanabad, Osmanabad 413501, Maharashtra, India*

A B S T R A C T

In this paper, we explore the feasibility of developing a comprehensive intrusion detection system tailored for the gas pipeline industry within contemporary artificial intelligence frameworks. This system aims to alert a gas controller to unexpected alterations in pipeline operational parameters, such as pressure, time intervals, delta pipeline PSI, and flow rate. Our investigation assesses the potential of employing artificial intelligence pattern recognition techniques, specifically utilizing Neural Network, for detecting gas system leaks, akin to the SCADA rate of change approach utilized in the hazardous liquids pipeline industry.

To achieve this, features were extracted from the dataset through the elimination of redundant information and data cleansing. A significant contribution of this research involves the utilization and the application of the neural network with three layers, each containing 25 units, 25 training rounds, and two layers with 60 training rounds, as demonstrated in the practical implementation of this work in Matlab. This application aims to identify and predict potential attacks in the gas pipeline industry.

The preliminary findings suggest promise in the field of artificial intelligence analysis, particularly in gas pipeline burst detection under the tested conditions. However, further exploration is required to transform this concept into an effective crack monitoring method, particularly through real-world assessments in complex system setups. Given the distinct physical behaviors of gases and liquids under varying pressure and flow conditions, the direct applicability of AI in predicting variations in Pipeline PSI and total delta pipeline PSI is challenging.

For instance, scenarios like by-passing, back-feeding, and other system-specific conditions demand customized solutions using AI. Consequently, collaborative efforts with AI modeling experts are essential for a more comprehensive understanding of the practicality of this automated technology adaptation.

Keywords: intrusion detection, artificial intelligence, pressure, time intervals, delta pipeline PSI, flow rate, Neural Network

1. Introduction

The primary objective of this research is to scrutinize historical rupture data obtained from gas transmission operators, assessing the viability and efficacy of applying this information to natural gas pipelines. The study aims to make recommendations based on the identification of potential attacks in the gas pipeline industry, employing automated means such as machine learning with pre-existing publicly available datasets. Furthermore, this endeavor reflects the industry's collaborative spirit in seeking a practical and efficiently managed solution for detecting natural gas attacks, with considerations for both cost and effort.

Detecting information leaks swiftly on natural gas pipelines poses challenges due to the compressible nature of natural gas and its varied transportation methods. While existing techniques like Real Time Transient Models (RTTM) can offer internal-based information leak detection on natural gas pipelines, configuring and maintaining these systems presents engineering challenges. Moreover, the dynamic nature of interconnected natural gas transmission networks may render RTTMs impractical for certain operations.

Conventional volume-balance Computational Pipeline Monitoring (CPM) systems, commonly applied in liquid pipelines, prove ineffective and prone to false alarms when applied to natural gas pipelines. This is attributed to the unique physical properties of natural gas under pressure. Alternatively, SCADA-based applications, requiring no hydraulic model, utilizing existing pipeline instrumentation, and requiring only minimal enhancements to traditional SCADA logic, offer a more promising approach for monitoring gas transmission pipelines for information leaks or potential external attacks.

The proposed methodology, grounded in well-known machine learning techniques, seeks to automate the process of intrusion detection in the gas pipeline industry. Leveraging SCADA systems, which commonly use Rate of Change (ROC) alarms to notify controllers of abrupt pressure and flow changes, this research suggests extending ROC alarm functionality to encompass a "pattern of alarms" concept. This composite alarm approach aims to identify

leak events and ruptures while minimizing false or nuisance alarms stemming from regular operational activities, thus enhancing the reliability and scalability of information recognition and response in a timely manner.

Implementing advanced alarm management practices is crucial for minimizing nuisance alarms and enhancing the relevance of alerts in process control systems. A recommended approach is the utilization of "pattern of alarms" techniques, which can replace irrelevant alarms with more meaningful ones, aligning better with the evaluated variables. The specific application under examination in this research project employs a pattern of rate of change alarms, termed "Rate of Change Combination," particularly effective in detecting ruptures and attacks in liquids pipelines.

This study aims to assess the applicability of this "Rate of Change Combination" technique in detecting attacks or intrusions in the gas pipeline industry, leveraging a publicly available dataset and various machine learning-based techniques. Utilizing these values individually as inputs for a simple detection method or combining any two of them as a composite attack prediction can lead to detection during normal operations. The distinctive signature of an intrusion or attack emerges when all three conditions occur within a short time span.

To configure multiple inputs for individual rate of change monitoring for potential attacks, the subsequent step involves assigning them to specific pipe segments or regions. When the rate of change conditions aligns with the defined logic for all three algorithms, the application generates a higher priority prediction. This indicates that all the specified conditions, signaling a potential information leak or attack, have been detected in the provided dataset.

Recognizing the challenges associated with complex and increasingly risky projects, Honeywell has transformed oil and gas automation and safety projects. Achieving capital savings of up to 30%, Honeywell's Lean automation engineering methodology involves parallel engineering, standard cabinet design, and the integration of three enabling technologies: Cloud Engineering, Universal Channel Technology, and Virtualization. These elements collectively contribute to taking automation off the critical project path, streamlining processes, and enhancing project efficiency.

2. Background

The confidential dataset [4] generously provided real-world data on pipeline ruptures and attacks, forming the basis of the evaluation for the machine learning application. The purpose was to collectively analyze four inputs, aiming to confidently identify a rupture or attack signature while minimizing false alarms. The request for data emphasized the importance of "related" pipeline inputs, ideally encompassing upstream and downstream pressures and flow rates where available. In instances where flow rate data was unavailable, multiple related pressure data points were used, represented by delta pipeline PSI. The provided data included a brief pipeline layout description, specifying the locations of inputs relative to the attack or rupture site (e.g., PS1 located approximately x miles upstream of the attack location). Additional requested pipeline information covered fundamental details such as pipe alignment, nominal operating pressure, and the method used to determine the leak or rupture.

It is crucial to clarify that the research's primary focus was on assessing the machine learning application's capability to detect attack conditions, not on evaluating the application's robustness in generating false positive alarms under specific operating conditions. Consequently, the collected data from the dataset pertained to specific attack incidents rather than scenarios that might trigger false positives under certain conditions, as mentioned in article [5].

The initial step involved identifying and configuring a robust "rate of change" monitor for each input. Traditional SCADA rate of change evaluations are often noisy, leading to numerous false alarms due to uncertainties in poll times, fast or interrogate scanning, and data latency. This study improved the reliability of rate of change evaluations by using a configurable number of samples instead of a fixed time, enhancing accuracy after data cleansing. While the inputs were designed for pressure and flow in attacking situations, the same algorithm could be applied to other inputs like control valve position or compressor rpm.

With the Rate of Change Combination record fully configured, SCADA polling of the points commenced through a simulated remote telemetry unit (RTU) that read the time-series data in real-time. This approach allowed the representation of live data in the machine learning application as it occurred during rupture events, facilitating the assessment of the application's effectiveness, as elucidated in article [5].

As the data was polled, each configured point was monitored by comparing the first sample's value against the last value, calculating the rate of change of the process by dividing this difference by the time span between them. These points, normally unseen during alarm violations, served the application's goal of eliminating individual alarms and focusing on predetermined combinations that activate machine learning algorithms.

Several methodologies employing various machine learning techniques have been explored as references for predicting intrusions in the gas pipeline industry as demonstrated in below table 1:

Method Used	Reference Details
K-nearest neighbor	(Wang, P.; Yu, B.; Han, D.; Li, J.; Sun, D.; Xiang, Y et al., 2018) [8]
Follow-The-Leader	(Rios-Mercado, R.; Borraz-Sanchez et al., 2015) [5]
Self-Organizing Map	(Gato, L.; Henriques, J et al., 2015) [7]
Adaptive Vector Quantization	(Chaczykowski, M & Sumaili Akilimali , 2010) [1]
Probabilistic Neural Network	(Gato, L.; Henriques, J et al., 2015) [7]
Fuzzy And Arima	(Pambour, K.A.; Bolado-Lavin, et al., 2016) (2016) [4]
Support Vector Machine	(Sundar, K.; Zlotnik, et al., 2018) [9]

Table 1. Demonstration of Methods used in Background in Gas pipeline Industry.

3. Course of Action

Equations The dataset in question encompasses 12 inputs covering a thirty-minute time interval, with updates occurring once per minute. This dataset captures normal operations, with seven minutes preceding an attack and 23 minutes following the attack. The provided pipeline PSI data corresponds to five interconnected main lines. The attack detection occurred when the pressure difference between two interconnected lines surpassed a configurable alarm threshold, utilizing conventional machine learning functionality.

The test configuration for the dataset involves setting up four selected pipeline inputs as machine learning algorithm records, intended for combination. Feature extraction and format conversion processes are outlined in Figure 1. The Matlab application enhances the reliability of rate of change evaluation by utilizing a configurable number of samples, as opposed to a fixed time, to assess the rate of change. This evaluation is instrumental in computing the accuracy of detection for the algorithms employed in this work.

Key parameters in the configuration include:

1. ROC Violation Endpoint:

- This represents the threshold that the rate of change must surpass, either positively or negatively, to establish a data cleaning space for different values from the dataset. The chosen value determines the region of consideration for attack detection.

2. ROC Suppression Endpoint:

- This value sets the limit that the rate of change must surpass, either positively or negatively, to suppress the ROC calculation. It influences the prediction accuracy of detection.

3. Sample Number:

- This parameter denotes the number of samples used for the ROC calculation, influencing the accuracy prediction of detection.

4. Exceed Limit Time Endpoint:

- This parameter determines the duration for which a ROC stays in violation or suppression once it enters that state. It is a crucial factor in understanding the temporal aspects of the detection process.

Figure 1. Step by step processing of feature extractions and final file creation for consumption.

The attack detection methodology relies on configuring these parameters effectively, ensuring that the rate of change analysis aligns with the characteristics of the dataset and the specific conditions associated with potential attacks on the gas pipeline. The MATLAB application, through its configurable settings, facilitates a nuanced evaluation of the rate of change for accurate and timely intrusion detection.

3.1 Neural Networks Approach:

Artificial Neural Network (ANN) is an artificial intelligence technique inspired by the structure and functioning of the human brain, designed to solve complex problems. It is a versatile tool applicable to both supervised and unsupervised classification problems, making it suitable for tasks such as predicting information leaks and potential gas pipeline attacks. ANN has found diverse applications, including bankruptcy prediction, fault detection, speech recognition, and product inspection [10].

The structure of an ANN consists of three types of nodes: input, hidden, and output nodes, as depicted in Figure 2. The input nodes receive the initial information and transmit signals to other nodes. In the context of gas pipeline security, multiple instances of input data are processed through the hidden layers, typically consisting of 2-3 layers, leading to the prediction of potential attacks or information leaks in the dataset. The output nodes gather information from the network nodes and produce the final output, which can be interpreted as a prediction of potential attacks or information leaks in the gas pipeline industry.

The three types of nodes play distinct roles:

1. Input Nodes:

- These nodes receive the input information, representing the features of the dataset related to the gas pipeline system. The input nodes emit signals to the hidden layers for further processing.

2. Hidden Nodes:

- Positioned between the input and output nodes, the hidden nodes process the incoming signals, capturing complex patterns and relationships within the data. Multiple hidden layers allow for the extraction of hierarchical features that contribute to the network's predictive capabilities.

3. Output Nodes:

- The output nodes receive information from the hidden layers and produce the final output, indicating the prediction of potential attacks or information leaks. This output is based on the patterns and relationships learned during the training phase of the ANN.

Notably, the hidden nodes neither directly interact with external sources nor receive information from the external environment. Their role is to transform the input signals to capture underlying patterns that contribute to the network's ability to predict potential attacks or information leaks in the gas pipeline industry. This structure enables ANNs to learn and generalize from the input data, making them powerful tools for solving complex classification problems.

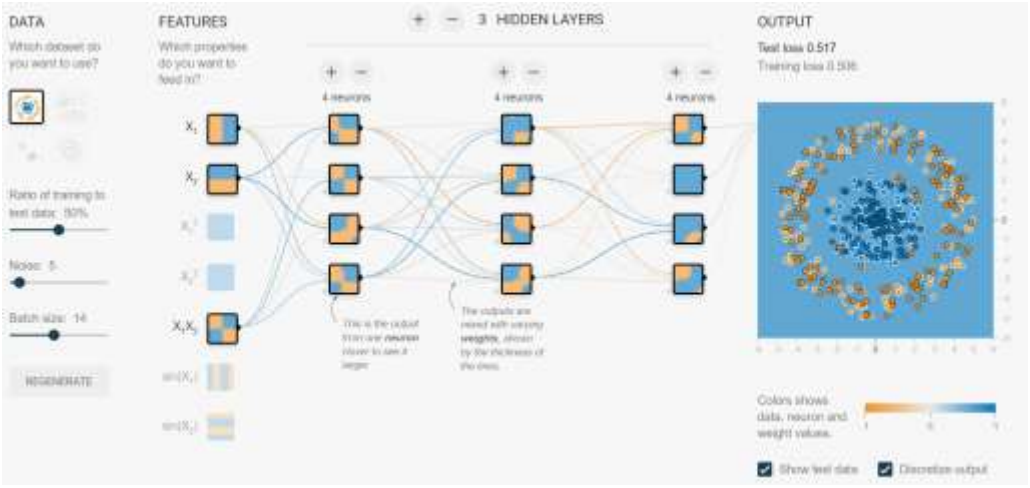

Figure 2: Demonstration of ANN with 3 hidden layers and multiple nodes processed through layers for prediction of the potential attacks.

4. Results

All

4.1 Neural Networks results:

The accuracy of detection, illustrated in Figure-3, is indicative of the intrusion detection (ID) performance concerning instances in the dataset. Notably, there was a temporal gap between the initiation of the data and the occurrence of the actual rupture, information leak, or any form of intrusion incident. Throughout this timeframe, the pipeline PSI was continually measured utilizing neural network methodology for flow measurement, aiming to identify potential intrusions within the dataset.

The neural network methodology employed in this context involved monitoring the pipeline PSI for flow measurement and assessing whether it exceeded a predefined violation limit. When this limit was surpassed, it triggered an internal prediction mechanism for variations in pipeline PSI, total delta pipeline PSI, and intrusion classification accuracy. The outcomes of this process are depicted in Figure-4.

Iteration	Model Prediction	Actual True	Accuracy
1	8	8	1
2	3	3	1
3	3	1	0
4	4	4	1
5	1	1	1
6	3	3	1
7	2	3	0
8	6	6	1
9	9	9	1
10	4	4	1
11	5	5	1
12	4	7	0
13	4	4	1
14	3	3	1
15	7	7	1
16	3	3	1
17	4	3	0
18	2	2	1
19	6	6	1
20	8	8	1
21	5	5	1
22	9	9	1
23	3	7	0
24	2	2	1
25	5	5	1
26	6	6	1
27	7	7	1
28	5	5	1
29	6	6	1

Figure 3. Demonstrations of Models prediction vs Actual prediction each iteration and stating its Accuracy.

Figure-4 showcases the results of applying neural network methodology to measure pipeline PSI, illustrating the variations in total delta pipeline PSI, and presenting the accuracy of intrusion classification. This approach leverages the capabilities of neural networks to recognize patterns and anomalies in the pipeline data, providing a means of early detection for potential incidents like ruptures or information leaks. The neural network methodology contributes to the proactive monitoring of the gas pipeline system, enhancing the overall security and reliability of operations.

PSI Value
~0.1792
Variation Value
~0.2954
Total PSI for Delta Pipeline
~0.1733
Gas Inspection Classification Accuracy
~0.9166

Figure 4. Demonstrating key performance parameter values from our proposed Neural Network Approach

5. Conclusion

Through the analysis of a dataset. The evaluation indicates that the neural network approach yields the highest accuracy among various historical approaches in predicting intrusions in the gas pipeline industry.

Upon analyzing the provided dataset, which includes sample data related to ruptures, information leaks, and intrusions in the gas pipeline industry, the Matlab application, incorporating three main machine learning algorithms, demonstrates the potential to predict intrusions in the probabilistic dataset. Further development of this application could enhance its capacity to recognize ruptures in natural gas pipelines. This aligns with the recommendation suggesting the implementation of an automatic SCADA system that trends data alarms, which would improve controller awareness of abnormal conditions such as pipeline ruptures, data leaks, and server memory loss. Subsequently, controllers could take appropriate actions to respond to these situations.

Beyond the analyzed Matlab application, other techniques have been developed and may warrant evaluation for natural gas leak detection. One such approach is the "alarm bracketing" or clamping method. This method enables gas controllers to activate an alarm bracket for a pipeline, grouping hydraulically related pressure inputs, pipeline PSI, and delta pipeline delta PSI. After feature extraction, the current pressures for all points in the bracket group are read, creating an operating envelope that reflects normal operating conditions. The intrusion classification is triggered when pressures and pipeline PSI deviate from this envelope without any operational reason. This offers a straightforward yet effective rupture monitoring, information leak, and intrusion detection application that employs common machine learning techniques by compressing the dataset for pressures, flows, and rate of change. The program logic is designed to accommodate multiple-point inputs for intrusion detection and prediction in the gas pipeline industry.

References

[1] Chaczykowski, M. Transient flow in natural gas pipeline—The effect of pipeline thermal model Adaptive Vector Quantization. Appl. Math. Model. 2010, 34, 1051–1067.

[2] Nguyen, H.; Chan, C. Optimal scheduling of gas pipeline operation using genetic algorithms. In Proceedings of the Canadian Conference on Electrical and Computer Engineering using Machine Learning, Saskatoon, SK, Canada, 1–4 May 2012.

[3] Zlotnik, A.; Chertkov, M.; Backhaus, S. Optimal control of transient flow in natural gas networks using machine learning. In Proceedings of the 54th IEEE Conference on Decision and Control, Osaka, Japan, 15–18 December 2015.

[4] Pambour, K.A.; Bolado-Lavin, R.; Dijkema, G.P.J. An integrated transient model for simulating the operation of natural gas transport systems. J. Nat. Gas Sci. Eng. 2016, 28, 672–690.

[5] Rios-Mercado, R.; Borraz-Sanchez, C. Optimization problems in natural gas transportation systems using Follow-The-Leader (Fdl): A state-of-the-art review. Appl. Energy 2015, 147, 536–555.

[6] Behrooz, H.; Boozarjomehry, R. Modeling and state estimation for gas transmission networks using Machine Learning Algorithms. J. Nat. Gas Sci. Eng. 2015, 22, 551–570.

[7] Gato, L.; Henriques, J. Dynamic behaviour of high-pressure natural-gas flow in pipelines using Probabilistic Neural Network and Self-Organizing Map. Int. J. Heat Fluid Flow 2015, 26, 817–825.

[8] Wang, P.; Yu, B.; Han, D.; Li, J.; Sun, D.; Xiang, Y.; Wang, L. Adaptive implicit finite difference method for natural gas pipeline transient flow. Oil Gas Sci. Technol using Neeural Network. Dec-2018.

[9] Sundar, K.; Zlotnik, A. State and parameter estimation for natural gas pipeline networks using transient state data. IEEE Trans. Control Syst. Technol. 2018, 99, 1–15.

[10] Durgut, I.; Leblebicioglu, K. Optimal control of gas pipelines via infinite-dimensional analysis. Int. J. Numer. Methods Fluids 2016, 22, 867–879.
[11] Cortinovis, A.; Mercangoz, M.; Zovadelli, M.; Pareschi, D.; de Marco, A.; Bittanti, S. Online performance tracking and load sharing optimization for parallel operation of gas compressors. Comput. Chem. Eng. 2016, 88, 145–156.

[12] Wen, K.; Xia, Z.; Yu, W.; Gong, J. A new lumped parameter model for natural gas pipelines in state space. Energies 11 June 2017.

[13] B. Durakovic, “Thermal Performances of Glazed Energy Storage Systems with Various Storage Materials: An Experimental study”, Sustainable Cities and Society, vol. 45, pp. 422-430, 2019.

[14] B. Durakovic, “Design for Additive Manufacturing: Benefits, Trends and Challenges”, Periodicals of Engineering and Natural Sciences (PEN), vol. 6, pp. 179–191, 2018.

[15] “Intrusion detection system in gas-pipeline industry using machine learning”, Ali Hasan Dakheel, Awfa Hasan Dakheel , Haider Hadi Abbas - http://pen.ius.edu.ba/index.php/pen/article/viewFile/512/371