

## UM ESTIMADOR DE LOCALIZAÇÃO DE VAZAMENTO EM DUTOS COM O ALGORITMO DE APRENDIZADO DE MÁQUINA SUPERVISIONADO K-NN

Thiago J. Gomes Oliveira<sup>1</sup>, João G. de Oliveira Lima<sup>2</sup>, Fernando A. Dantas Gomes Pinto<sup>3</sup>

1. Estudante de Sistemas de Informação do Centro Universitário CESMAC.

2. Centro Universitário CESMAC / Orientador

3. LaCCAN – Laboratório de Computação Científica e Análise Numérica (UFAL)

### Resumo:

Vários tipos de fluidos são transportados em dutos no Brasil e no Mundo, desde água, gases diversos e até o petróleo. Esse transporte traz vantagens competitivas assim como riscos às operações. Com o tempo, estes dutos se desgastam provocando vazamentos no ambiente em que eles se encontram. Este problema fez com que diversas técnicas de detecção fossem criadas, cada uma com suas vantagens e desvantagens.

Através de buscas bibliográficas, viu-se a necessidade de implantar um método de baixo custo operacional e de baixa especialização e que fosse capaz de estimar a localização do vazamento. O objetivo desta pesquisa foi desenvolver um algoritmo capaz de estimar, através de uma série histórica de dados coletados por sensores de pressão em dutos, a localização de vazamentos com a utilização de classificador K-NN.

**Palavras-chave:** programação; pressão; inteligência artificial.

**Apoio financeiro:** Centro Universitário CESMAC.

### Introdução:

Entre as soluções energéticas disponíveis no mundo, o petróleo é o principal componente da matriz energética mundial, pois apresenta baixo custo, se comparado às demais alternativas, e uma ampla gama de utilização em diversos setores (REIS, 2011). Com uma capacidade de exploração estimada em 2,609 milhões de barris de petróleo por dia (PORTAL BRASIL), o Brasil ocupa a 15ª posição (BOL, 2016) entre os países que o exploram. Esta grande capacidade produtiva traz consigo responsabilidades socioambientais nas diversas etapas da cadeia produtiva e uma das áreas que vem recebendo muita atenção é a malha de dutos. Para escoar esta produção, a Transpetro, maior empresa de logística da América Latina, é responsável por mais de 14mil km de oleodutos e gasodutos em todo Brasil (TRANSPETRO).

O problema atribuído a estes dutos está relacionado às suas estruturas que são fixadas e instaladas no subsolo de regiões densamente povoadas. Com o tempo, estes dutos sofrem corrosão em suas paredes, o que pode provocar a perda de fluidos para o ambiente externo. O risco de vazamento sobre estas regiões faz com que órgãos competentes adotem rígidas normas de prevenção a acidentes com riscos socioambientais. Segundo (ANP), um relatório foi produzido para garantir procedimentos técnicos quanto aos testes de prevenção a acidentes nestes dutos.

Diante dos riscos que podem provocar esta perda, hoje são investidos muitos recursos em técnicas e métodos de detecção de vazamentos que tentam minimizar eventuais danos ambientais. O trabalho de (COLOMBAROLI, 2009) apresenta métodos que tradicionalmente mais se destacam. Todos estes possuem a capacidade de detectar e localizar vazamentos.

Outros fluidos transportados em dutos como água potável e gases dos mais diversos tipos causam perdas financeiras e danos ambientais quando acontecem vazamentos em suas redes.

Esta pesquisa objetivou investigar um novo sistema baseado em modelagem computacional, de baixo custo operacional e de baixa especialização, que utiliza técnicas de aprendizagem de máquina para auxiliar na detecção e, principalmente, na localização de vazamentos em dutos. Para este trabalho, experimentos preliminares foram realizados (PINTO, 2011) com o auxílio da ferramenta Weka da Universidade de Waikato, Nova Zelândia. Ela é uma ferramenta de Descoberta de Conhecimento em Banco de Dados que contempla uma série de algoritmos de pré-processamento de dados, de aprendizagem de máquina e de avaliação de resultados.

### Metodologia:

Para esta pesquisa foi escolhido o método de mudança de pressão, porém trazendo uma nova abordagem, utilizando Arduino como plataforma para o desenvolvimento. Este é uma plataforma de prototipagem eletrônica e utiliza como linguagem de programação C/C++. O Arduino chamou atenção na sua praticidade e facilidade, não exigindo tanto conhecimento em eletrônica, além de ser acessivo. Como por exemplo, uma placa custa de R\$ 50,00 a R\$ 100,00 e o seu ambiente de desenvolvimento é gratuito e Open Source (ARDUINO), visando o baixo custo operacional e a baixa especialização.

Foi criado um ambiente de simulação para ser feita a coleta dos dados, composto de um cano PVC de 2m (metros) de comprimento e 32mm (milímetros) de diâmetro representando um duto, 2 sensores de pressão Mpx5700ap, 3 válvulas para simular as rupturas de fluido gasoso.

O experimento consistia em colocar ar dentro do cano para que os sensores conseguissem fazer a leitura da variação da pressão e no momento em que uma das válvulas fosse aberta, simulando a ruptura,

estes sensores captariam a mudança na pressão, lida pela placa do Arduino e apresentada no monitor do computador, mostrando a queda. Para cada ruptura, os dados coletados fizeram parte de uma série histórica dos valores de pressão interna na tubulação para cada área específica (pré-definida).

O modelo computacional proposto utilizou como técnica de aprendizagem de máquina o algoritmo *K-Nearest Neighbors*. Este é um classificador supervisionado que resolve problemas de classificação (valores de saída discretos) ou regressão (valores de saída contínuos) para novos elementos a um conjunto de treinamento que não possuem classes definidas. Esta técnica procura por *K* elementos (*k*-vizinhos mais próximos) de um conjunto de treinamento que esteja mais próximo deste elemento desconhecido (que tenha a menor Distância Euclidiana). Verificando-se quais são as classes destes *k*-vizinhos e a classe mais frequente será atribuída a classe do novo elemento (PINTO, 2011).

Com os dados coletados, esta série histórica das rupturas foi submetida a um algoritmo de classificação escrito na linguagem de programação Java 8 com uso da API *Weka* para K-NN (*weka.classifiers.lazy.IBk*), em uma máquina MacBook Pro, com 8 GB de memória RAM e com capacidade de armazenamento de 1 TB de dados. A IDE (*Integrated Development Environment*) para o desenvolvimento do algoritmo e análise dos dados foi o *Eclipse Mars (Release 4.5.2)*.

### Resultados e Discussão:

O estudo desenvolvido em um ambiente simulado e controlado, apresentou resultados condizentes com a hipótese levantada, trazendo uma perspectiva positiva quanto à continuação das pesquisas para o desenvolvimento de um simulador maior.

Como resultados deste experimento, os sensores conseguiram ler a queda de pressão de maneira diferente em cada quadrante. Quando a válvula do primeiro quadrante foi aberta, o primeiro sensor sentiu a queda da pressão mais rápido que o segundo. No momento em que a válvula do segundo quadrante foi aberta, os dois sensores sentiram a queda quase ao mesmo tempo. Já no terceiro quadrante, ao abrir a válvula, o segundo sensor sentiu a queda da pressão mais rápido que o primeiro. Também foram feitas leituras para verificar a pressão sem nenhuma válvula aberta, simulando o duto sem nenhuma ruptura. A unidade de medida que se utilizou neste experimento foi kPa (quilopascal).

Abaixo temos uma tabela com alguns dados colhidos dos dois sensores, representando o valor da pressão no cano sem nenhuma ruptura e no momento que ocorre a ruptura em um quadrante específico.

RUPTURA NO QUADRANTE 1		RUPTURA NO QUADRANTE 2		RUPTURA NO QUADRANTE 3		SEM RUPTURA	
SENSOR1 (kPa)	SENSOR2 (kPa)	SENSOR1 (kPa)	SENSOR2 (kPa)	SENSOR1 (kPa)	SENSOR2 (kPa)	SENSOR1 (kPa)	SENSOR2 (kPa)
2,01	1,99	1,99	1,99	1,99	1,97	0,67	0,67
1,98	1,99	1,98	2,00	1,99	1,95	1,24	1,23
1,96	1,98	1,97	1,97	1,98	1,94	1,82	1,80
1,92	1,97	1,94	1,94	1,96	1,88	2,02	2,02
1,90	1,95	1,94	1,95	1,92	1,84	2,14	2,16
1,91	1,91	1,92	1,91	1,90	1,84	2,17	2,17
1,90	1,92	1,91	1,91	1,87	1,83	2,19	2,16
1,86	1,89	1,90	1,90	1,46	1,43	2,19	2,21
1,83	1,87	1,87	1,88	1,13	1,12		
1,46	1,49	1,79	1,79	0,80	0,77		
1,13	1,32	1,46	1,45	0,80	0,75		
0,93	0,99	1,13	1,13	0,79	0,76		
0,80	0,86	0,80	0,81	0,78	0,74		
0,79	0,81	0,72	0,71	0,74	0,71		
0,74	0,74	0,68	0,68	0,72	0,70		
0,72	0,72	0,67	0,66	0,71	0,70		

Após a coleta, todos os dados foram colocados em um arquivo .arff com a formatação necessária para serem utilizados como base de treinamento na Weka.

Esta ferramenta é uma coleção de algoritmos vindos de diferentes áreas da inteligência artificial dedicada ao estudo de aprendizagem de máquina. Sua base tem a análise computacional e a estatística dos dados fornecidos por intermédio de técnicas de mineração de dados tentando, indutivamente, gerar hipóteses para soluções a partir dos padrões encontrados. Os algoritmos podem ser aplicados diretamente a um conjunto de dados ou chamados dentro do próprio código Java. Ela também contém instrumentos para o pré-processamento de dados, classificação, regressão, agrupamento, regras de associação e visualização (WEKA).

Ao selecionar este arquivo, foi escolhido o classificador K-NN, com  $K = 3$ . O valor de  $K$  representa a quantidade de vizinhos mais próximos de um dado elemento.

Ao rodar a ferramenta, tiveram-se os seguintes dados:

1. De 186 instâncias (dados coletados), 136 foram classificadas corretamente (73,1183%) e 50 incorretamente (26,8817%);

- A estatística Kappa foi 0,6238, onde os valores desta estatística variam de 0 (péssimo) a 1 (excelente);
- A matriz de confusão foi:

CLASSE	A	B	C	D
QUADRANTE1 = A	22	12	2	2
QUADRANTE2 = B	15	14	3	5
QUADRANTE3 = C	6	2	28	2
SEMFURO = D	1	0	0	72

- Para os 38 dados do QUADRANTE1, a ferramenta previu 22 para o QUADRANTE1, 12 para o QUADRANTE2, 2 para o QUADRANTE3 e 2 para SEMFURO.
- Para os 37 dados do QUADRANTE2, a ferramenta previu 15 para o QUADRANTE1, 14 para o QUADRANTE2, 3 para o QUADRANTE3 e 5 para SEMFURO.
- Para os 38 dados do QUADRANTE3, a ferramenta previu 6 para o QUADRANTE1, 2 para o QUADRANTE2, 28 para o QUADRANTE3 e 2 para SEMFURO.
- Para os 73 dados do SEMFURO, a ferramenta previu 1 para o QUADRANTE1, 0 para o QUADRANTE2, 0 para o QUADRANTE3 e 72 para SEMFURO.

Por fim, foi realizado um teste com 5 dados aleatórios retirados da base de treinamento. Estes dados serviram como entrada para que o algoritmo pudesse classificá-los de acordo com os testes realizados em cima da base de treinamento. Da mesma forma que o primeiro teste, o classificador escolhido foi o K-NN com o valor de K=3.

Os dados obtidos foram os seguintes:

- De 6 instâncias (dados), 6 foram classificadas corretamente (100%);
- A estatística Kappa foi 1;
- A matriz de confusão foi:

CLASSE	A	B	C	D
QUADRANTE1 = A	1	0	0	0
QUADRANTE2 = B	0	1	0	0
QUADRANTE3 = C	0	0	1	0
SEMFURO = D	0	0	0	3

- Para o dado do QUADRANTE1, a ferramenta previu 1 para o QUADRANTE1, 0 para o QUADRANTE2, 0 para o QUADRANTE3 e 0 para SEMFURO.
- Para o dado do QUADRANTE2, a ferramenta previu 0 para o QUADRANTE1, 1 para o QUADRANTE2, 0 para o QUADRANTE3 e 0 para SEMFURO.
- Para o dado do QUADRANTE3, a ferramenta previu 0 para o QUADRANTE1, 0 para o QUADRANTE2, 1 para o QUADRANTE3 e 0 para SEMFURO.
- Para os 3 dados do SEMFURO, a ferramenta previu 0 para o QUADRANTE1, 0 para o QUADRANTE2, 0 para o QUADRANTE3 e 3 para SEMFURO.

### Conclusões:

A simulação desenvolvida em um ambiente controlado, utilizando-se como instrumento de coleta de dados, dois sensores de pressão e um leitor de dados arduino, em um cano de PVC que serviu de base para a simulação de vazamento de fluido gasoso com perda de pressão no cano.

O levantamento feito permitiu inicialmente captar variações de pressão nos quadrantes marcado para simular o vazamento.

Levando-se em conta o que foi observado neste relatório, foi possível, através dos experimentos realizados com os dados tirados do ambiente de simulação e aplicados na Weka, estimar a localização do vazamento em um duto por meio de técnicas de aprendizagem de máquina com o classificador supervisionado K-NN.

O modelo K-NN conseguiu identificar e validar os dados por quadrante sendo necessário novos testes para ampliar a base de dados de variação de pressão, assim como criar ambientes de teste baseados em canos de diâmetros diferentes e com uma variação maior de furos.

O objetivo do trabalho foi alcançado no ambiente simulado e o modelo de supervisionado K-NN se mostrou eficiente para a medição de variação de pressão e assim conseguir identificar onde fica o fato gerador da variação de pressão, limitando assim a área de busca do vazamento.

Novos trabalhos e produtos podem ser gerados a partir dos resultados desta pesquisa, ficando assim um modelo barato que pode ter seu potencial expandido em novas pesquisas aplicadas em parceria com

empresas.

### Referências bibliográficas

ANP. **Regulamento Técnico de Dutos Terrestres para Movimentação de Petróleo, Derivados e Gás Natural** (RTDT), 2010.

BOL. **Os 15 países com as maiores reservas de petróleo**. Disponível em: <https://noticias.bol.uol.com.br/bolistas/os-15-paises-com-as-maiores-reservas-de-petroleo.htm> Acesso em 05 de março de 2017.

COLOMBAROLI, P. L. S. **Deteção de Vazamentos em Dutos de Petróleo**. In: 5º Congresso Brasileiro de P & D em Petróleo e Gás, 2009, Fortaleza - CE. Congresso Brasileiro de P & D em Petróleo e Gás, 2009. v. 5.

PINTO, F. A. D. G. **Avaliação dos Métodos de Aprendizado de Máquina Supervisionado Aplicados a Base de Dados Pima Indians Diabetes com a Ferramenta Weka**. Programa de Pós-Graduação em Engenharia de Software (Lato Sensu), Faculdade Alagoana de Tecnologia, Maceió - AL. 2011.

PORTAL BRASIL. **Produção de petróleo bate recorde pelo 3º mês consecutivo**. Disponível em: <http://www.brasil.gov.br/economia-e-emprego/2016/10/producao-de-petroleo-bate-recorde-pelo-3-mes-consecutivo> Acesso em 01 de março de 2017.

REIS, L. B. dos. **Geração de Energia Elétrica**. 2. ed. São Paulo: Manole, 2011.

TRANSPETRO. **Dutos e Terminais**. Disponível em: [http://www.transpetro.com.br/pt\\_br/areas-de-negocios/terminais-e-oleodutos.html](http://www.transpetro.com.br/pt_br/areas-de-negocios/terminais-e-oleodutos.html) Acesso em 01 de março de 2017.

TRANSPETRO. **Gás Natural**. Disponível em: [http://www.transpetro.com.br/pt\\_br/areas-de-negocios/gas-natural.html](http://www.transpetro.com.br/pt_br/areas-de-negocios/gas-natural.html) Acesso em 01 de março de 2017.

Site Oficial. **Weka**. Disponível em: <http://www.cs.waikato.ac.nz/ml/weka/> Acesso em 03 de junho de 2017.