UDC 60:620:004.9:006

Big Data Approach Application for Steel Pipelines in the Conditions of Corrosion Fatigue

Skrynkovskyy R. M.[1], Yuzevych L. V.[2,3], Ogirko O. I.[4], Pawlowski G.[5*]

[1] Lviv University of Business and Law, 99 Kulparkivska St., 79021 Lviv, Ukraine;
[2] Lviv Polytechnic National University, 12 Stepana Bandery St., 79013 Lviv, Ukraine;
[3] Karpenko Physico-Mechanical Institute of the NAS of Ukraine, 5 Naukova St., 79060 Lviv, Ukraine;
[4] Lviv State University of Internal Affairs, 26 Horodotska St., 79007 Lviv, Ukraine;
[5] Zaklad Handlowo-Uslugowy BHP, 17 Kostrzynska St., 69-113 Gorzyca, Poland

Abstract. This paper presents results of the use of Big Data approach and neural network for the pipelines diagnosis problem. In this case the pipeline is in the conditions of crack growth of corrosion fatigue and exposed to hydrogen. It is proposed to use graphene protective coatings. The mathematical model for estimating the changes in the effective surface energy of WPL during plastic deformation, electrochemical overstrain, polarization potential and current density of the metal dissolution reaction at the top of the crack on the pipeline surface during its mechanical loading in an aqueous electrolyte solution is given. The dissolution of the metal is considered on the juvenile surface, taking into account the anode and cathode regions based on the approaches of surface physics and electrochemistry. An element of a mathematical model is a quality functional, taking into account information flows and a sensitivity coefficient. Functional quality is used to specify the feedback between the investment project methodology and risk estimates, as well as to optimize the information flows of enterprises and improve the system of protection of metallic underground pipelines that operate under conditions of corrosion fatigue. The purpose of this project is to improve the relevant regulatory and technical documents as well as software.

Keywords: gas pipeline, monitoring, fatigue crack, corrosion, databases, Big Data, neural network, intelligent software, hardware, databases.

1 Introduction

The problem of unstructured data is related to the separation of sources, their format and quality [1, 2].

Underground metal gas pipelines that come in contact with soil electrolyte can be taken as an example. Thermomechanical processing allows the yield strengths of pipe steels to be tailored through combinations of grain refinement, precipitation hardening (micro-alloying) and phase transformations [3].

Places of maximum stresses are the tips of cracks in the metal. Near the tips of cracks appears the influence of hydrogen. The influence of hydrogen leads to hydrogen embrittlement. That can be dangerous for metal pipelines.

Once hydrogen has been absorbed into the steel at the crack tip, the mechanism(s) responsible for material damage resulting from electrochemical and gaseous charging will be similar with most of the experimentally observed differences resulting from differences in the thermodynamics and kinetics of the dissociation reactions influencing the activity of the atomic hydrogen in the crack tip process zone [3].

Control of pipelines and technical equipment should take the standards, other regulatory documents, criteria of strength and reliability into account. In result, researchers will receive large amounts of unstructured data (Big Data).

It is necessary to create new methods for analyzing flows of information and chemical components for correct organizing, integrating and processing large data of underground piping systems. This is a problem of research.

2 Literature Review

Experienced pipeline operators utilize Magnetic Flux Leakage (MFL) sensors to probe oil and gas pipelines for the purpose of localizing and sizing different defect types [4]. A large number of sensors is usually used to cover the targeted pipelines. The sensors are equally distributed

around the circumference of the pipeline; and every three millimeters the sensors measure MFL signals [4]. Thus, the collected raw data is so big that it makes the pipeline probing process difficult, exhausting and error-prone. Machine learning approaches such as neural networks have made it possible to effectively manage the complexity pertaining to big data and learn their intrinsic properties [4]. Discriminant features, which characterize different defect depth patterns, are first obtained from the raw data. Neural networks are then trained using these features. The Levenberg-Marquardt back-propagation learning algorithm is adopted in the training process, during which the weight and bias parameters of the networks are tuned to optimize their performances [4].

The real-time text processing pipeline using open-source big data tools which minimize the latency to process data streams, explain it and evaluate is proposed in paper [5]. Proposed data processing pipeline on Apache Kafka for data ingestion, Apache Spark for in-memory data processing, Apache Cassandra for storing processed results, and D3 JavaScript library for visualization is new technology [5]. Apache Kafka is a distributed data transfer system that allows to process large amounts of data in real-time. The effectiveness of the proposed pipeline under varying deployment scenarios to perform sentiment analysis using Twitter dataset is evaluated [5].

A novel model for reasoning across components of Big Data Pipelines in a probabilistically well-founded manner is proposed [6]. The interaction of components as dependencies on an underlying graphical model is presented. Different message passing schemes on this graphical model provide various inference algorithms to trade-off end-to-end performance and computational cost. The framework with an efficient beam search algorithm is instantiated. That demonstrates its efficiency on two Big Data Pipelines: parsing and relation extraction [6].

An infrastructure for parallel analyzing big data in order to search, pattern recognition and decision-making based on the use of the Boolean metric of cyberspace measurement is proposed [7]. It is characterized by using only logical operation for determining the cyber-distance by means of cyclic closing at least one object, which allows significantly increasing the speed of analysis of large data [7].

Formulation of the research goals

The purpose of this study is to organize and integrate informative resources of Big Data in the system of a gas pipeline and cathodic protection device and corresponding technological, physical and chemical processes.

System of objects and processes

The objects of a gas transport system (GTS): metallic pipes; metallic and dielectric coverages; devices of cathode defence; devices of anodic defence; compressors; devices of diagnostics and control are considered. Additional objects of research: standards and other normative documents; software; technological, physical and chemical processes for the enumerated objects and processes characteristic large amount of information (Big Data) are taken into consideration.

A problem of the Big Data obtained during diagnosis of underground gas pipelines (UGP) and devices by means of contactless current measurements (CCM) is raised [8].

3 Research Methodology

3.1 Mathematical modeling of pipelines in the conditions of corrosive fatigue

The presence of fatigue cracks on the surface of metallic underground pipelines emphasise the problem of calculation values of strength characteristics at the action of corrosive environments that did not find the complete decision nowadays. In this connection it is necessary to correct the row of defects and normative and technical documents related to insufficient actuality of corresponding.

In normative documents from exploitation of construction elements on this time the reasonable norms of legitimate values of corrosive damages, reduction of bearing strength of construction elements are absent. It creates complications in the ground of normative terms of exploitation and evaluation of the maximum state of metallic constructions, in planning of charges on exploitation of construction elements and repair and restoration work.

An object of researches is underground metallic pipelines that are in the conditions of corrosion-fatigue destruction. The subject of a study is normative document. It is expedient to specify and perfect on the basis of the information acquiring of results monitoring of underground metallic pipelines functioning.

It is important to formulate the criterion and scientifically reasonable recommendations for providing of quality of underground metallic pipelines exploitation in the conditions of fatigue and influence of aggressive environment, and also forming of normative principles in sphere of pipeline transport.

Only underground main gas pipelines in the ground electrolyte in the conditions of low cycle fatigue should be considered. For the improvement of normative documents it is expedient to build a complex mathematical model, which will unite the physical and chemical model of corrosion-fatigue processes, model of piling up description of damages in metals and theory of risks elements.

Equation that binds length of fatigue corrosive crack and amount of cycles of loading to the coefficient of intensity of mechanical tensions is used for the modeling of speed increasing of fatigue crack on the middle rectilinear area of kinetic curve. Corresponding equation is improved for a metal in corrosive environment and the pH-value of environment. The electrode potential of metal is also taken into account.

For the base model of damage accumulation for metals in the conditions of irregular deformation based on the curvilinear change of damages model and energetic characteristics are used near description of process of low cycle fatigue. It is a base on criterion of fatigue strength.

Critical specific work that answers the origin of fatigue crack is included in a criterion.

Equation, power descriptions of metal and function of relative value of amplitude intensity of tensions, that characterizes the degree of mechanism influence of fatigue on the fatigue curve is also used for description of the irregular cyclic loading of metal. Evolutional equation is written in for the modeling of low cycle fatigue of metals.

Correlation for description of low cycle corrosion fatigue of material in metallic underground pipelines is complemented by equations for the evaluation of risks within the limits of investment project that executes corresponding organization (enterprise).

The functional of quality is used for receiving a feedback in methodology of risk evaluation of investment project and for optimization of informative streams of enterprise and improvement of the defence system of metallic underground pipelines from a corrosive fatigue. The aim of that application is the improvement of the corresponding normative and technical providing and software.

In zones with non-stationary plasticity strain it is expedient to use the criteria of adhesion strength, biocorrosive aggressiveness of soils, mechanical criterion for the coefficient of intensity of tensions (the overstrain of corrosive process takes into account), the criterion of corrosive stability of pitting, criterion correlation for the evaluation of speed of stability corrosion of metal in the defect of isolating coverage together with entered by diagnostic weight of signs and diagnostic value of inspections, that will complement, specify and perfect the system of the corrosive monitoring of pipelines and be used for control of corrosive process. With their help described and regulated by a state standard optimization of terms of construction elements defence of oil and gas industry can be conducted.

As a result a new complex mathematical model in relation to upgrading of corrosion protection of metallic underground pipelines from positions of corrosion fatigue, electrochemistry, physics of surface processes, mechanics of destruction and theory of risks is offered. The conducted modeling takes piling up of damages in metals into account and allows to study the mechanisms of distribution of corrosive fatigue cracks in underground metallic pipelines that are in aggressive environments, in particular, in saltwater and ground electrolyte. The results of mathematical modeling are the basis of methodology development and improvement of normative and technical documents for metallic underground pipelines, that are under the action of the regular and irregular cyclic loading in the conditions of low cycle corrosive fatigue.

The joint use of corrosive fatigue criteria and corrosive monitoring criteria of pipelines offered in this article will allow to study in detail the mechanisms of distribution of corrosive fatigue cracks in underground metallic pipelines that are in aggressive environments from positions of corrosive fatigue, electrochemistry, physics of surface processes, mechanics of fracture and theory of risks.

For optimizing information flows P_k the functional of quality $J(P_k, FB(P_k))$ with calculation for sensitivity coefficient β is used [9]:

$$J(P_k, FB(P_k)) = \int_{t_0}^{t_k} f(\bar{y}, \bar{u}, \bar{s}, \beta) dt \Rightarrow opt, \quad (1)$$

where $\bar{y}$ – vector of specific impacts ($y_j(t)$ – vector components (key parameters for the GTS), $j = 1, 2, \ldots, n$); $\bar{u}$ – control vector of information flows; $\bar{s}$ – vector of indeterminate perturbations; P_k – information flows for the GTS and security system ($k = 1, 2, \ldots, m$); m – total number of information flows P_k considered in the given GTS; $[t_0, t_k]$ – time interval, in which the process is considered (formation of optimal values of parameters corresponding to P_k; $f(\bar{y}, \bar{u}, \bar{s}, \beta)$ – function reflecting quality index; β – sensitivity coefficient; $FB(P_k)$ – function characterizing inverse relationship between flows P_k and project's environment with accounting for sensitivity coefficient β and expert opinions; opt – optimization symbol; t – time.

3.2 GTS protection system and Big Data

To protect the system (GTS) and related software, we recommend using the scientific work algorithm [10].

This algorithm (stages) allows the following [10]:

1. The data owner describes which data user will grant the access to certain data under specified constraints, and generates a policy rule, then sends the rule to the trusted authority.

2. The data owner encrypts the data with the encryption key, and then stores encrypted data in the database.

3. The encrypted data is sent into the Kafka cluster which is comprised of one or more servers each of which is called a broker.

4. The data consumer sends a request to the trusted authority for data access, which involves passing on the sticky policy.

5. The trusted authority checks policies, potentially including challenges to the data user.

6. If all the policy checks are fulfilled and validated, the trusted authority releases the private decryption key to the data consumer.

7. The data consumer can get the encrypted data from the kafka and decrypt it by using the decryption key.

3.3 Prospect of pipelines defence from corrosion with the help of graphene coverages

Graphene layers as corrosion resistant films are presented [10]. Mild steel coupons were coated from the synthesized graphene solution. Three layers of graphene films were able to reduce the corrosion rate by 99 % [11].

Further work needs to be done to test the durability of the graphene films and its resistance to corrosion with respect to time [12].

Besides having a unique mechanical and electronic properties [12], graphene coating has broad prospects for practical application in specific structural and functional materials [12, 13].

4 Results

An object of the research is the metal under mechanical stress with a surface crack in aqueous electrolyte solution. Destruction of passive films, a juvenile surface (JS) with width δ and zone of plastic deformation occurs at the crack tip under the action of stress [14]. A geometrical parameter δ at the crack tip in a first approximation is interpreted as its opening δ_{IC}. The crack tip and, in particular, JS spreads into the depth of body under mechanical stress and corrosive environment. The cathodic and anodic electrochemical reactions occur at the crack tip region. Corrosive dissolution corresponds to the anodic reaction of metal. The crack tip (JS) is interpreted as an anode (A), beyond it on sides it is interpreted as a cathode area (K) [14]. The system "A − K" presents an electrochemical pair.

On the basis of the correlations [14] coefficient of stress intensity (KSI) K_{1SCC} is related to crack opening δ_{IC} and overstrain ζ of reaction of dissolution of metal by next formulas (2):

$$K_{1SCC} = \sqrt{\frac{E}{1-v^2} \cdot \left(WPL - z_{si} F \rho \delta \frac{\zeta}{M} \right)} = \sqrt{E \cdot \sigma_T \cdot \delta_{1C}}, \quad (2)$$

where z_{si} – microcrack, m; M – is molecular mass of metal, g/·mol^{-1}; K_{1SCC} – is a threshold value KSI, that is minimum value that corresponds to the beginning of corrosion crack propagation; WPL – is specific energy spending on the plastic deformation of surface layer of body during formation of new (juvenile) surface; E, v – is the modulus of longitudinal elasticity (Young's modulus) of material is a formal charge of the solvated ions; F = 96 500 C·mol^{-1} – Faraday constant; δ – is an impending front of and Poisson's ratio.

An electrochemical overstrain is a deviation of electrode potential from its equilibrium (in relation to solution near the electrode) thermodynamics value during polarization of electrode under electric current [14].

The empiric correlation that binds KSI to WPL is set in article [14] based on research results of pin (contact) deformation of different brands of steel (Poland: 15GA, EU: 20MoCr2-2, USA: A 516-55, etc.):

$$K_{1SCC} = a_1 \sqrt{WPL} - a_2;$$
$$a_1 = 2.26 \cdot 10^8 \frac{\sqrt{N}}{m}; \quad (3)$$
$$a_2 = 6.98 \sqrt{m} \ (MPa).$$

Correlation (Kaesche) for the current density i_a in the crack tip according to the paper [14] is:

$$i_a = \frac{\alpha \cdot \chi \cdot \Delta \psi_{ak}}{\delta \cdot \ln(c/\delta)}, \quad (4)$$

where α – is an angle of the crack tip; χ – is conductivity of electrolyte; $\Delta \psi_{ak}$ – is a change of potential between anodic and cathodic parts; c – is a depth of crack. Expression (4) is submitted for a crack in a unstressed metal. In actual use of construction elements, in particular pipelines, it is necessary to take diagnostic method and the terms of corrosion stress into account [14, 15]. Therefore correlation (4) should be generalized by addition of information about mechanical parameters and characteristics.

Correlations (1)–(4) present a mathematical model for change estimating of effective surface energy WPL during a plastic deformation, electrochemical overstrain and density of current of reaction of dissolution of metal in the crack tip on the surface of metal at its mechanical stress in aqueous electrolyte solution. Such type of research is required by the uses of Big Data.

For example we will use the experimental data that approximate dependence as Tafel correlation for the evaluation of influence of mechanical stress tension on intensity of corrosive processes in Steel 20, that is in 3 % solution of NaCl, in particular, in the crack tip in the moment of fracture of passive films, when the anodic current of i_a grows substantially [12]:

$$i_a = i_0 \exp(DE/a), \quad DE = E_0 - E_a, \quad (5)$$

where i_0 is the corrosive current; a is a Tafel parameter of anodic process; E_0, E_a is the corrosion potential and potential for anodic process. The polarization potential E_P of the metal surface (pipeline) in [8] is presented.

We will use the correlation:

$$E_P = f(DE, E_0, E_a, \zeta). \quad (6)$$

A mathematical model (1)–(6) is developed for the evaluation of surface energy of plastic deformation, overstrain, polarization potential and density of current of metal dissolution reaction in the crack tip for the metal (steel) loaded in aqueous electrolyte solution on the basis of approaches of surface physics and electrochemistry. Dissolution of metal is considered on a juvenile surface taking into account a stress intensity coefficient (KSI).

We propose to use the information of Big Data in the process of neural network spectral analysis, which is able to adapt to requirements of a specific sensor [16]. Such component features with high reliability, ability to adapt to a specific application, and usage of design principles ensuring the possibility of a simple expansion of the intelligent component potential by means of completion with new algorithmic solutions [16].

Let's formulate general information on data collection (Big Data) regarding pipelines and their processing:

1. Analysis of processes.
2. Modeling.
3. Monitoring using device БВС-К [8].
4. Data mining.
5. Optimization of information.
6. Optimization of development processes.
7. Assessment of production risks for pipelines.
8. Statistical methods of information analysis.
9. Application of methods of machine learning.
10. Obtaining stable estimates of model parameters.
11. Construction of prediction and an estimation of a resource of pipelines on the basis of the learned models.
12. Decision-making regarding repair terms.

5 Conclusions

A mathematical model (1)–(6) is developed for the evaluation of surface energy of plastic deformation, over-strain, polarization potential and density of current of metal dissolution reaction in the crack tip for the metal (steel) loaded in aqueous electrolyte solution on the basis of approaches of surface physics and electrochemistry. Dissolution of metal is considered on a juvenile surface taking into account a stress intensity coefficient.

The general principles for the selection of information concerning the monitoring of underground pipelines based on Big Data technology as a result of data processing and the corresponding algorithm are formulated.

A method of functioning of intelligent software and hardware complex for the monitoring system of a metal gas pipelines and security system using Big Data is proposed.

In this paper the Big Data methodology is improved due to functional of quality application, micro and macro processes and inverse relationships.

References

1. Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, Vol. 36, No. 4, pp. 1165–1188.

2. Pavlyshenko, B. M. (2016). Linear, machine learning and probabilistic approaches for time series analysis. *IEEE International Conference on Data Stream Mining and Processing (DSMP)*, Aug. 2016. doi: 10.1109/dsmp.2016.7583582.

3. Nanninga, N., Slifka, A., Levy, Y., & White, C. (2010). A review of fatigue crack growth for pipeline steels exposed to hydrogen. *Journal of Research of the National Institute of Standards and Technology*, Vol. 115, No. 6, pp. 437, doi: 10.6028/jres.115.030.

4. Mohamed, A., Hamdi, M. S., & Tahar, S. (2015). A Machine Learning Approach for Big Data in Oil and Gas Pipelines. *3rd International Conference on Future Internet of Things and Cloud*.

5. Nazeer, H., Iqbal, W., Bokhari, F., Bukhari, F., & Baig, S. (2017). Real-time Text Analytics Pipeline Using Open-source Big Data Tools. *Distributed, Parallel, and Cluster Computing*, pp. 1–6.

6. K. Raman, A. Swaminathan, J. Gehrke, and T. Joachims, "Beyond myopic inference in big data pipelines," Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '13, 2013. doi: https://doi.org/10.1145/2487575.2487588 .

7. Hahanov, V. I., Litvinova, E. I., Chumachenko, S. V., Yemelianov, I., & Amer, T. B. (2016). Processor structures for Big Data analysis. *Radio Electronic and Computer Systems*, No. 6(80), pp. 163–175.

8. Dzhala, R. M., Verbenets', B. Y., Mel'nyk, M. I., Mytsyk, A. B., Savula, R. S., & Semenyuk, O. M. (2017). New Methods for the Corrosion Monitoring of Underground Pipelines According to the Measurements of Currents and Potentials. *Materials Science*, Vol. 52, No. 5, pp. 732–741.

9. Krap, N., & Yuzevych, V. (2013). Neural Networks as a tool for managing the configurations of tourist flow projects. *Management of Development of Complex Systems*, No. 14, pp. 37–40.

10. Thein, K. M., Nyunt, T. T., & Aye, K. N. (2017). Security of real-time Big Data analytics pipeline. *International Journal of Advances in Electronics and Computer Science*, Vol. 4, No. 2, pp. 1–5.

11. Pavan, A. S. S., & Ramanan, S. R. (2016). A study on corrosion resistant graphene films on low alloy steel. *Applied Nanoscience*, Vol. 6, No. 8, pp. 1175–1181, doi: 10.1007/s13204-016-0530-2.

12. Koman, B. P., & Yuzevich, V. M. (2015). Energy Parameters of Interfacial Layers in Composite Systems: Graphene – (Si, Cu, Fe, Co, Au, Ag, Al, Ru, Hf, Pb) and Semiconductor (Si, Ge) – (Fe, Co, Cu, Al, Au, Cr, W, Pb). *Journal of Nano- and Electronic Physics*, Vol. 7, No 4, pp. 04059-1–04059-7.

13. Novoselov, K. S., Morozov, S. V., Mohinddin, T. M. G., Ponomarenko, L. A., Elias, D. C., Yang, R., Barbolina, I. I., Blake, P., Booth, T. J., Jiang, D., Giesbers, J., Hill, E. W., & Geim, A. K. (2007). Electronic properties of grapheme. *Physica Status Solidi*, Vol. 244, No. 11, pp. 4106–4111, doi: 10.1002/pssb.200776208.

14. Yuzevych, V. M., Dzhala, R. M., & Koman, B. P. (2018). Analysis of Metal Corrosion under Conditions of Mechanical Impacts and Aggressive Environments. *Metallofizika i Noveishie Tekhnologii*, Vol. 39, No. 12, pp. 1655–1667, doi: 10.15407/mfint.39.12.1655.

15. Yuzevych, V., Klyuvak, O., & Skrynkovskyy, R. (2016). Diagnostics of the system of interaction between the government and business in terms of public e-procurement. *Economic Annals-XXI*, Vol. 160, No. 7–8, pp. 39–44, doi: 10.21003/ea.v160-08.

16. Cielen, D., Meysman, A. D. B., & Ali, M. (2016). *Introducing Data Science: Big Data, Machine Learning, and More, using Python Tools*. Manning Publications, New York, USA.

Застосування підходу **Big Data** для сталевих трубопроводів в умовах корозійної втоми

Скриньковський Р. М.[1], Юзевич Л. В.[2, 3], Огірко О. І.[4], Павловські Г.[5]

[1] Львівський університет бізнесу та права, вул. Кульпарківська, 99, м. Львів, 79021, Україна;
[2] Національний університет "Львівська політехніка", вул. Степана Бандери, 12, м. Львів, 79013, Україна;
[3] Фізико-механічний інститут ім. Г. В. Карпенка НАН України, вул. Наукова, 5, м. Львів, 79060, Україна;
[4] Львівський державний університет внутрішніх справ, вул. Городоцька, 26, м. Львів, 79007, Україна;
[5] Компанія "Zaklad Handlowo-Uslugowy BHP", вул. Костшинська, 17, м. Гужиця, 69-113, Польща

Анотація. У роботі подано результати використання підходів Big Data та нейронних мереж для діагностування трубопроводів. Приймаємо до уваги, що на поверхні трубопроводу знаходяться втомні корозійні тріщини і метал піддається впливу водню. Запропоновано використовувати графенові захисні покриття. Наведено елементи математичної моделі для оцінювання змін ефективної поверхневої енергії WPL під час пластичної деформації, електрохімічного перенапруження, поляризаційного потенціалу та густини струму реакції розчинення металу у вершині тріщини на поверхні трубопроводу під час його механічного навантаження у водному розчині електроліту. Розчинення металу розглядаємо на ювенільній поверхні з урахуванням анодної та катодної ділянок на основі підходів фізики поверхні та електрохімії. Елементом математичної моделі є функціонал якості з урахуванням інформаційних потоків та коефіцієнта чутливості. Функціонал якості використовуємо для конкретизації зворотного зв'язку між методологією інвестиційного проекту та оцінками ризику, а також для оптимізації інформаційних потоків підприємств та вдосконалення системи захисту металевих підземних трубопроводів, які функціонують в умовах корозійної втоми. Метою цього проекту є вдосконалення відповідних нормативних та технічних документів, а також програмного забезпечення.

Ключові слова: газопровід, моніторинг, тріщина втоми, корозія, бази даних, великі дані, нейронна мережа, інтелектуальне програмне забезпечення, апаратні засоби, бази даних.