Matrix model of data and knowledge presentation to revealing regularities of the fluid flow regime in a pipeline based on hydrodynamics parameters

A. Yankovskaya[1,2,3,4], A. Travkov[2]

[1]Tomsk State University of Architecture and Building, 634003, Tomsk, Russia
[2]National Research Tomsk State University, 634050, Tomsk, Russia
[3]National Research Tomsk Polytechnic University, 634050, Tomsk, Russia
[4]Tomsk State University of Control Systems and Radioelectronics, 634050, Tomsk, Russia

Abstract

The study offers an original solution to one of the problems of hydrodynamics, namely revealing the regularities in the flow regime of fluid in a pipeline depending on the hydrodynamic parameters. The solution is based on using the intelligent system of the regularities revealing and decision-making. For the first time, a matrix model of data and knowledge representation (MM) is used for these purposes in the form of two matrices: descriptions of the fluid state in the space of characteristic features of hydrodynamics (pressure, velocity, temperature, and others); its rows are associated with various combinations of characteristic features values, and distinguishing of the diagnostic type, whose rows are associated with the corresponding rows of the description matrix, and its columns are associated with the two classifying features. The first classifying features takes four values corresponding to four fluid flow regimes, and the second classifying features takes three values only for the turbulent flow regime value from the first classifying features.

Keywords: matrix model; description matrix; distinguishing matrix; data and knowledge representation; regularities; fluid flow regimes; hydrodynamics; intelligent system; regularities revealing; decision-making

1. Introduction

Development of computer systems for current and precise determination of the fluid flow regime in a pipeline [1] is an extremely urgent issue in exploiting pipelines [2-3]. If a flow regime is determined inaccurately, it may interfere with or stop production, lead to breakdowns or other undesirable consequences. Such consequences may result in considerable expenses of finance and time. However, studies of the existing methods for calculating the fluid flow regime in a pipeline [3-7] have shown that these methods do not fully include all the parameters (features), which affect the fluid flow in a pipeline. This issue is of great importance to control fluidic flow inside oil pipes [8] because transportation of hydrocarbons is very dangerous process that requires continuous monitoring.

It is well known fact that every flow regime can be described with a set of parameters (features) determining the behavior of the fluidic flow. The exact type of the fluid flow depends on a number of features, such as velocity, viscosity, density, pressure, and more.

Taking into consideration all these factors, especially a large number of parameters influencing the type of fluidic flow in pipeline, it becomes clear that modern computer systems should be used to study regularities between these hydraulic parameters and support diagnostic decisions on the fluidic flow regimes. For these purposes, it is obviously rational to create intelligent systems (IS) for diagnostics of the fluid flow regimes in a pipeline (IS DFFRP) to determine various regularities, and also decision-making and their justification for such diagnostics.

Unlike the modeling methods described in the papers [10-11], proposed IS DFFRP is being developed specifically for oil transportation pipelines.

The Intelligent Systems Laboratory of Tomsk State University of Architecture and Building (TSUAB) under A.E. Yankovskaya's supervision has achieved significant results in the field. The researchers have developed 3 original intelligent instrumental software (IIS) for constructing applied intelligent systems. More than 30 applied intelligent systems are based on them. These systems are intended for different areas, such as geology, geoecology, medicine, ecology, electronics, psychology and other, more than 30 applications overall [12–25]. These IIS and applied intelligent systems are based on test methods of pattern recognition intended for various regularities revealing and decision-making, and also for making and justification decisions using cognitive tools. Thus, it is logical to apply the developed IIS to create an IS DFFRP.

The next section describes the matrix model to represent data and knowledge of hydraulics, their structuring, fragment of the description and distinguishing matrices as well as directions of the future research.

2. Matrix model of data and knowledge representation. Regularities

The IS DFFRP currently being developed is based on the matrix model of data and knowledge representation [25].

The matrix model includes an integer matrix of descriptions **Q** and matrix of distinguishing **R** [25]. They represent a learning sample with objects belonging to the known (determined by experts) patterns.

The rows of the **Q** matrix are corresponded with the s_i learning objects ($i = \overline{1, N}$, where N is the quantity of objects), columns are corresponded with the z_j characteristic features ($j = \overline{1, M}$, where M is the quantity of features, which together represent the description of each object). The element $q_{i,j}$ of the **Q** matrix sets the value of the jth feature for the ith object.

The data and knowledge base is formed on the base of the matrix model of data and knowledge representation [25], which includes the description integer matrix (**Q**) setting descriptions of objects in the space of k-valued $z_1, ..., z_m$ characteristic features, and the distinguishing integer matrix (**R**), which sets the partition of objects into equivalence classes according to each mechanisms of the classification. If the value of a feature is not essential for the object, it is marked with a dash ("−") in the respective element of the **Q** matrix. For each z_j ($j \in \{1,2,..., m\}$) feature, either an interval of its value change, or an integer value is set.

Rows of the **R** matrix are associated with the rows of the **Q** matrix, columns are associated with the k_j classification features ($j = \overline{1, L}$, where L is the quantity of classification mechanisms partitioning learning objects into equivalence classes). An $r_{i,j}$ element of the **R** matrix sets belonging of an ith object to a certain class (by designation its number) based on the jth classification mechanism.

Objects with same combination of the classification features k_j, corresponding to a certain final solution, belong to the same pattern. This means that a number of patterns are equal to a number of non-repeating rows of the matrix **R** and equal as well to the subset of the rows from the matrix **Q** assigned to the same rows from the matrix **R.** These mutually assigned rows describe patterns.

It should be noted that proposed model allows us to represent not only data but also experts' knowledge, because one row of the matrix **Q** describes subset of the similar objects in the form of an interval (using a dash "−"). All these objects have the same final solution set by the corresponding row from the matrix **R**.

Let us consider that the objects from the learning sample do not contain of measuring and entry errors. Otherwise it is necessary to reveal data and knowledge inconsistencies using a special subsystem realized in IS DFFPR.

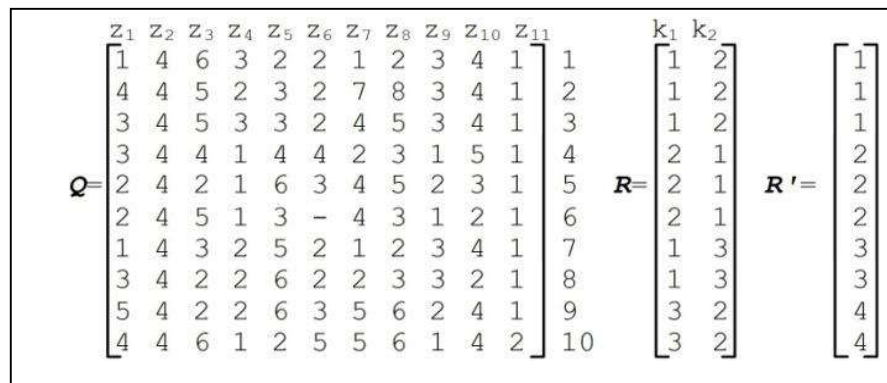Figure 1 shows an example of matrix representation of data and knowledge.

Fig. 1. Example of a description **Q** matrix and a distinguishing **R** and **R'** matrix.

One of the important means of the data and knowledge analysis [25] are diagnostic tests, i.e. tests distinguishing objects from different patterns [25] constructed during of the regularities revealing in the data and knowledge base, which are represented by the **Q**, **R** matrices. These tests are used for decision-making in the IS based on the methods of test pattern recognition.

Regularities are subsets of features with particular, easy-to-interpret properties that affect the distinguishing ability of objects from different patterns that are stably observed for objects from the learning sample and are exhibited in other objects of the same nature and weight coefficients of features that characterize their individual contribution [25] to the distinguish ability of objects and the information weight given, unlike, on the subset of tests used for a final decision-making.

These subsets include constant (taking the same value for all patterns), stable (constant inside a pattern, but nonconstant), non–informative (not distinguishing any pair of objects), alternative (in the sense of their inclusion in DT), dependant (in the sense of the inclusion of subsets of distinguishable pairs of objects), unessential (not included in any irredundant DT), obligatory (included in all irredundant DT), and pseudo-obligatory (which are not obligatory, but included in all IUDT involved in decision-making) features, as well as all minimal and all (or part, for a large feature space) irredundant distinguishing subsets of features that are essentially minimal and irredundant DTs, respectively. The weight coefficients of characteristic features are also included in regularities [25], as well as the information weight of characteristic features.

Regularities of the fluid flow regime in a pipeline include the described regularities used for decision-making and their justification. The revealed regularities will allow a significant decrease the number of measurements to determine the fluid flow regime in a pipeline.

3. Data and knowledge structuring. An illustrative example of presenting data and knowledge on the fluid flow regimes in a pipeline

3.1. Data and knowledge structuring

According to the described matrix model of data and knowledge representation, for the purpose of to determine the fluid flow regime in a pipeline, data and knowledge in hydraulics were structured.

Since the limits of the paper, it is impossible to describe the whole feature space, whose dimension exceeds 30, so only a part of it and a fragment of the description and distinguishing matrices are introduced.

Based on the analysis of a range of papers [3-11, 26-29], 9 real characteristic features and their values used in forming the **Q** matrix were formulated. As mentioned above, real characteristic features are represented by intervals of their values. Characteristic features and partitioning intervals for each of the 9 characteristic features represented by integers, are listed in the Table 1.

Table 1. The list of the characteristic features and their partitioning intervals.

Characteristic features	Code	Value intervals
Flow velocity (m/s)	z_1	1 – up to 0.5 inclusive; 2 – from 0.5 to 1; 3 – from 1 to 1.5; 4 – from 1.5 to 2; 5 – from 2 to 2.5; 6 – from 2.5 to 3; 7 – from 3 to 3.5 ; 8 – from 3.5 to 4; 9 – from 4 to 4.5; 10 – from 4.5 to 5; 11 – from 5 to 5.5; 12 – from 5.5 to 6; 13 – from 6 to 6.5; 14 – from 6.5 to 7;
Viscosity (mPa·s)	z_2	1 – from 0.2 to 0.5; 2 – from 0.5 to 0.8; 3 – from 0.8 to 1; 4 – from 1 to 1.2; 5 – from 1.2 to 1,5; 6 – from 1,5 to 1.8; 7 – from 1.8 to 2 ; 8 – from 2 to 2.5; 9 – from 2.5 to 3; 10 – from 3 to 4; 11 – from 4 to 5; 12 – from 5 to 8; 13 – from 8 to 10; 14 – from 10 to 20; 15 – over 20;
Density (kg/m³)	z_3	1 – from 550 to 580; 2 – from 580 to 610; 3 – from 610 to 640; 4 – from 640 to 670; 5 – from 670 to 700; 6 – from 700 to 730; 7 – from 730 to 760 ; 8 – from 760 to 790; 9 – from 790 to 820; 10 – from 820 to 850; 11 – from 850 to 880; 12 – from 880 to 910; 13 – from 910 to 940; 14 – from 940 to 970; 15 – from 970 to 1000;
Cross-sectional area (m²)	z_4	1 – up to 0.3 inclusive; 2 – from 0.3 to 0.5; 3 – from 0.5 to 0.7; 4 – from 0.7 to 0.9; 5 – from 1 to 1.2; 6 – from 1.2 to 1.4;
Element of hydraulic power unit	z_5	1 – circular pipe (smooth); 2 – flexible tubing; 3 – smooth concentric annulus; 4 – tap valve; 5 – dispensable slide-valve port; 6 – plate and poppet valves; 7 – strainer;
Temperature (°C)	z_6	1 – from 0 to 10; 2 – from 10 to 20; 3 – from 10 to 20; 4 – from 20 to 30; 5 – from 30 to 40; 6 – from 40 to 50; 7 – from 50 to 60; 8 – from 60 to 70; 9 – from 70 to 80; 10 – from 80 to 90; 11 – from 90 to 100;
Type of fluid	z_7	1 – water; 2 – sea water; 3 – oil; 4 – ether; 5 – alcohol; 6 – gasoline; 7 – kerosene;
Roughness (mm)	z_8	1 – 0.0001; 2 – 0.001; 3 – 0.006; 4 – 0.015; 5 – 0.017; 6 – 0.02; 7 – 0.025; 8 – 0.1; 9 – 0.15; 10 – 0,25;
Pressure (MPa)	z_9	1 – from 0.25 to 0.75; 2 – from 0.75 to 2.5; 3 – from 2.5 to 5; 4 – from 5 to 6.4;

It should be noted, that the lines of the **Q** matrix are associated with the fluid flow regimes and represent only a part of various combinations of characteristic features values.

Table 2 contains the classification features and their values for the R matrix.

Table 2. The list of classification features and their values.

Characteristic features	Code	Values
Fluid flow regime	k_1	1 – ideal; 2 – laminar; 3 – transient; 4 – turbulent;
Zones of turbulent fluid flow regime	k_2	1 – zone of hydraulically smooth pipes; 2 – zone of mixed friction; 3 – zone of square-law resistance.

3.2. An illustrating example of representing data and knowledge about the fluid flow regimes in a pipeline

Figure 2 gives an illustrating example of data and knowledge representing on the fluid flow regimes in a pipeline. The illustrating example is a fragment of matrix description of data and knowledge in hydraulics.

The description matrix (Fig. 2) contains 9 columns associated with the mentioned above characteristic features, and 15 rows filled out by us.

The rows of the **R** matrix are associated with the rows of the **Q** matrix, the columns are associated with the aforementioned characteristic features k_j ($j \in \{1,2\}$). The $r_{i,j}$ element of the distinguishing matrix sets the belonging of the ith object (fluid flow) to a class based on the jth classification mechanism (fluid flow regime) by way of indicated the class number. A row of the **R** matrix sets the fluid flow regime and the zone of the turbulent regime for the studied field of hydraulics.

As mentioned above, the set of all non-repeating rows of the **R** matrix is associated with the set of selected patterns represented by the single-column **R'** matrix, whose elements are numbers of the patterns. This model does not permit intersecting objects from different patterns. Presence of such intersections is revealed via analysis using IS DFFRP.

Taking into consideration that with the limits of the paper do not allow us to fully present the matrix description of data and knowledge, Fig. 2 contains only a fragment of matrix data and knowledge representation on fluid flow regimes in a pipeline (**Q, R,** and **R'** matrices). The aforementioned fragment, which represents partial (only a part of characteristic feature space and its

values are used to determine the fluid flow regime) description of knowledge represented in the **Q** matrix containing 9 columns and 15 rows, while the **R** matrix – 2 columns.

$$
\begin{array}{ccccccccc}
z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 & z_8 & z_9
\end{array}
$$

$$
Q =
\begin{bmatrix}
1 & 5 & 15 & 2 & 1 & 5 & 1 & 1 & 1 \\
1 & 11 & 1 & 2 & 1 & 7 & 3 & 1 & 1 \\
1 & 13 & 1 & 2 & 1 & 8 & 3 & 1 & 1 \\
2 & 12 & 1 & 2 & 1 & 6 & 3 & 1 & 1 \\
2 & 4 & 1 & 4 & 1 & 4 & 3 & 6 & 5 \\
2 & 2 & 2 & 4 & 1 & 4 & 3 & 6 & 5 \\
2 & 2 & 6 & 1 & 1 & 4 & 3 & 6 & 1 \\
2 & 2 & 5 & 1 & 1 & 4 & 3 & 1 & 1 \\
3 & 2 & 5 & 1 & 1 & 3 & 3 & 1 & 1 \\
3 & 1 & 5 & 1 & 1 & 6 & 3 & 1 & 1 \\
2 & 12 & 1 & 2 & 1 & 7 & 3 & 1 & 1 \\
1 & 13 & 1 & 2 & 1 & 8 & 3 & 1 & 1 \\
2 & 12 & 1 & 2 & 1 & 4 & 3 & 1 & 1 \\
3 & 3 & 1 & 3 & 1 & 6 & 3 & 1 & 1 \\
2 & 12 & 1 & 2 & 1 & 5 & 3 & 1 & 1
\end{bmatrix}
\quad
R =
\begin{bmatrix}
k_1 & k_2 \\
4 & 1 \\
2 & - \\
2 & - \\
3 & - \\
4 & 3 \\
4 & 3 \\
4 & 3 \\
4 & 1 \\
4 & 2 \\
4 & 2 \\
3 & - \\
3 & - \\
2 & - \\
4 & 1 \\
3 & -
\end{bmatrix}
\quad
R' =
\begin{bmatrix}
4 \\
2 \\
2 \\
3 \\
6 \\
6 \\
6 \\
4 \\
5 \\
5 \\
3 \\
3 \\
2 \\
4 \\
3
\end{bmatrix}
\begin{array}{c}
1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 11 \\ 12 \\ 13 \\ 14 \\ 15
\end{array}
$$

Fig. 2. Fragment of the description **Q** matrix and the distinguishing **R** and **R'** matrix.

The fragment of the description and distinguishing matrices includes only 3 types of fluid flow regimes in the pipeline: laminar, transient, turbulent, and 3 aforementioned zones only for the turbulent regime.

The complete description and distinguishing matrices will be represented in the data and knowledge base of the IS DFFRP based on the fluid flow regimes in the pipeline, and the IS DFFRP is destined for prompt determination of various combinations of fluid flow regimes with the zones of the turbulent regime: 1) ideal flow regime; 2) laminar flow regime; 3) transient fluid flow regime; 4) turbulent flow regime, zone of hydraulically smooth pipes; 5) turbulent flow regime, zone of mixed friction; 6) turbulent flow regime, zone of square-law resistance. These 6 regimes are represented in the **R'** matrix.

4. Conclusion

Based on the performed analysis of modern-day state of research in determination of fluid flow regimes in a pipeline, for the first time it is proposed to create an intelligent system for diagnostics of the fluid flow regime in a pipeline designed to determine various regularities in parameters of hydrodynamics, which affect the fluid flow regimes, and also for decision-making and its justification on diagnostics of the fluid flow regime in a pipeline.

We also showed advisability of applying the matrix model for data and knowledge representation in the intelligent system for diagnostics of the fluid flow regime in a pipeline. The system is developed by us. For the first time, in accordance with the suggested model of representing data and knowledge, the feature space was formed; data and knowledge in the studied field were structured; characteristic and classification features were determined; real values of features were recoded into integer ones; the illustrating example representing partial description of data and knowledge in a matrix form was given. Herewith, in order to determine the fluid flow regime, we used only a part of the characteristic feature space and its values.

The IS DFFRP, which is being developed, will enable prompt and cheaper determination of the fluid flow regime in a pipeline, based on the hydrodynamic parameters under study: ideal; laminar; transient; turbulent, and also the zones of the turbulent regime: hydraulically smooth pipes; mixed friction; square-law resistance. A company serving the pipeline will obtain the ability to operative reaction to any changes in a pipeline and take the appropriate measures of managing the flow and eliminating the possibility of breakdowns.

The IS DFFRP, which we are developing, may be applied on direct appointment at various types of enterprises, for scientific purposes, as well as to train specialists in hydrodynamics.

Acknowledgment

Reference

[1] SNiP 2.01.07-85 «Loads and Impacts ». M.: Gosstroy of the USSR, 1985; 48 p. (in Russian)
[2] SNiP 2.05.06-85 «Trunk pipelines». M.: Gosstroy of Russia, 2001; 86 p. (in Russian)
[3] Brandt I. Multiphase Flow. Euroil, 1991: 62–63
[4] Roache PJ. Computational Fluid Dynamics. M.: Mir, 1980; 616 p. (in Russian)
[5] Yakovlev PV, Hodzhamuradova NB, Guba OE. Reynolds's analogy in heatmass exchange model at the turbulent mode of the convective movement of liquid in limited volume. Vestnik ASTU 2008; 6(47): 75–77. (in Russian)
[6] Nikonova VT, Sautkina TN. Determination of coefficient of hydraulic friction upon transition from the laminar mode of the movement of liquid to turbulent. Improvement of methods of hydraulic calculations of water throughput and treatment facilities 2009; 1(35): 65–68. (in Russian)

[7] Sargsyan NM. About the modes of the movement of viscous liquid in horizontally located round pipe under isothermal conditions. Chemical technology 2013; 2: 104–112. (in Russian)

[8] Kudinov VI. Bases of oil and gas business. M.: IKI, 2005; 720 p. (in Russian)

[9] Shammazov AM, Bakhtizin RN, Mastobayeva BN, Soshchenko AE. Pipeline transport of Russia. Pipeline transport of oil 2001; 2: 406 p. (in Russian)

[10] Edwards DA, Gunasekera D, Morris J, Shaw G, Shaw K, Walsh D, Fjerstad PA, Kikani J, Franco J, Hoang V, Quettier L. Reservoir simulation: Keeping Pace with Oilfield Complexity. Oilfield Review 2012; 5(23): 4–15.

[11] Aziz I, Brandt I, Gunasekera D, Hatveit D, Havre K, Weisz G, Xu ZG, Nas S, Spilling KE, Yokote R, Song S. Multiphase flow simulation-optimization field productivity. Oilfield Review 2015; 1(27): 26–37.

[12] Yankovskaya AE, Gedike AI, Ametov RV, Bleikher AM. IMSLOG-2002 Software Tool for Supporting Information Technologies of Test Pattern Recognition. Pattern Recognition and Image Analysis 2003; 13: 4: 650–657.

[13] Yankovskaya AE, Il'inskikh NN. On the Question of the Development and Application of Intelligent Biomedical Systems. Pattern Recognition and Image Analysis 1998; 8(3): 470–472.

[14] Yankovskaya AE. An Automaton Model, Fuzzy Logic, and Means of Cognitive Graphics in the Solution of Forecast Problems. Pattern Recognition and Image Analysis 1998; 8(2): 154–156.

[15] Rocher G, Brattstrum A, Gho S, Francenson F, Hintricus H, Mauser R, Packianather M, Pogrzeba G, Yankovskaya A, Zvegintsev V. Real-time Recognition of ECG by Using Powerful Information and Communication Technology for Intelligent Monitoring of Risk Patients. Application, Trends, Vcisions, VDE World Microtechnologies (MICRO.tech). Proceedings of International Congress. Hannover, Germany, 2000; 2: 759–762.

[16] Ryumkin A, Yankovskaya A. Intelligent Expansion of the Geoinformation System. The 6th German-Russian Workshop "Pattern Recognition and Image Understanding" OGRW-6-2003. Workshop proceedings. Novosibirsk, Russia, 2003; 202–205.

[17] Yankovskaya AE, Chernogoryuk GE, Muratova EA. Intelligent Test Recognizing Biomedical System. The 6th German-Russian Workshop "Pattern Recognition and Image Understanding" OGRW-6-2003. Workshop proceedings. Novosibirsk, Russia, 2003; 248–251.

[18] Yankovskaya AE, Semenov ME. Intelligent system for knowledge estimation on the base of mixed diagnostic tests and elements of fuzzy logic. Proceedings of the IASTED International Conference Technology for Education. December 14-16, Dallas, USA, 2011; 108–113.

[19] Yankovskaya AE, Kitler S. Mental Disorder Diagnostic System Based on Logical-Combinatorial Methods of Pattern Recognition. Computer Science Journal of Moldova 2013; 21(3-63): 391–400.

[20] Yankovskaya AE, Yamshanov A. Bases of intelligent system creation of decision-making support on road-climatic zoning. Pattern Recognition and Information Processing (PRIP'2014): Proceedings of the 12th International Conference (28–30 May 2014, Minsk, Belarus). Minsk : UIIP NASB, 2014; 311–315.

[21] Yankovskaya AE, Efimenko S, Cherepanov D. Structurization of data and knowledge for the information technology of road-climatic zoning. Applied Mechanics and Materials 2014; 682: 561–568.

[22] Yankovskaya AE, Dementyev Y, Lyapunov D, Yamshanov A. Intelligent Information Technology in Education. Information Technologies in Science, Management, Social Sphere and Medicine (ITSMSSM-2016). Atlantis Press Publishing, 2016: 17–21.

[23] Yankovskaya AE, Gorbunov I, Hodashinsky I, Chernogoryuk G. On a Question of the Information Technology Construction Based on Self-learning Medicine Intelligent System. Information Technologies in Science, Management, Social Sphere and Medicine (ITSMSSM-2016). Atlantis Press Publishing, 2016: 22–28.

[24] Yankovskaya AE, Shelupanov AA, Mironova VG. Construction of Hybrid Intelligent System of Express-Diagnostics of Information Security Attackers Based on the Synergy of Several Sciences and Scientific Directions. Pattern Recognition and Image Analysis 2016; 26(3): 524–532.

[25] Yankovskaya AE. Logical Tests and Cognitive Graphics Means in Intelligent System. Proc. 3rd All–Russian Conf. with International Participation "New Information Technologies in Discrete Structures Research" (Izd. SO RAN, Tomsk), 2000: 163–168. (in Russian)

[26] Birkhoff G. Hydrodynamics: A study in logic, fact, and similitude. M.: Foreign literature, 1954; 184 p. (in Russian)

[27] Loytsyansky LG. Mechanics of liquid and gas. M.-L.: Gostekhizdat, 1950; 479 p. (in Russian)

[28] Ayala LF, Adewumi MA. Low–Liquid Loading Multiphase Flow in Natural Gas Pipelines. Journal of Energy Resources and Technology 2003; 4(125): 284–293.

[29] Li C, Liu E, Yang Y, Najafi M, Ma B. The simulation of steady Flow in Condensate Gas Pipeline. Advances and Experiences with Pipelines and Trenchless Technology for Water, Sewer, Gas, and Oil Applications. Reston, Virginia, USA, American Society of Civil Engineers, 2009: 733–743.