Features of the synthesis of information and measurement systems using machine learning for conducting of environmental monitoring

Yu M Avdeev¹,₆, A I Pakhomova², A L Zolkin³, V D Munister⁴ and V M Kalyakina⁵

¹ Department of Urban Cadastre and Geodesy of the Engineering and Construction Institute, Vologda State University, Lenin Street 15, Vologda, 160000, Russia
² Department of Economics and management, Institute of service and entrepreneurship (branch) of DSTU in Shakhty, Shevchenko Street 147, Shahti, Rostov region 346500, Russia
³ Computer and Information Sciences Department, Povolzhskiy State University of Telecommunications and Informatics, L.Tolstogo Street 23, Samara 443010, Russia; Natural Sciences Department, Private institution of higher education "Medical University" Reaviz, Chapaevskaya Street 227, Samara 443001, Russia
⁴ Department of Mathematical Methods and Computer-Aided Design, Private educational institution "Donetsk Academy of Transport", Dzerzhinsky Street 7, Donetsk 83086, Ukraine; Department of Enterprise Economics, State educational institution of higher professional education "Donetsk National Technical University", Shibankova Street 2, Pokrovsk, Donetsk region 85300, Ukraine
⁵ Department of health, safety and environmental engineering protection, Don State technical University, Gagarin square 1, Rostov-on-Don 344002, Russia

E-mail: avdeevyur@yandex.ru

Abstract. The article is dedicated to the consideration of new methods for the synthesis of digital environmental monitoring systems, inclusive to the analysis through the methodology of machine learning. The role of circuit components proposed for implementation is considered. Attention is paid to the specifics of the analysis of already measured data and to the memory device of information-measuring systems. The importance of using MS-triggers and logistic regression in analysis of the stability of sensors is justified.

1. Rationale
An essential sign of modernity is the adaptation of existing technical mechanisms for assessment, management, and measurement to the general outline. It is necessary to determine the impact of digital transformation of the economy, giving a specific meaning for the presented object of study [1].

First of all, it shall be noted that of time and other costs in the entire time management of the industrial sector are minimized through the use of innovative mechanisms for selection, ranking, planning, and a new decision-making paradigm.
Secondly, the emphasis on the less pronounced industrialization, hence, more non-commercial character of the measuring component shall be identified. This is dictated by the concern of modern society about the real state of the ecological situation in the world.

2. Some aspects of the output values of digital environmental monitoring systems in the post-industrial era

The approach expressed through the so-called ticket quota system of rewards and punishments, both at the international level and at the domestic level (in which the idea of monitoring systems as a mandatory component for large industries has developed) is replaced by the development of similar systems for private persons [2].

For previous twenty years it was a niche component, an analogue of a technical hobby known in a narrow circle of enthusiasts who are fond of telemetry and processing. But the model of the Internet of Everything (or its particular case, better known as the Internet of Things) demonstrates the autochthonous nature of distribution of the principle of measurements and their processing.

An important drawback of modern information-measuring systems is the insufficient degree of reduction of functional requirements to edictal metrological metrology [3]. It is explained be the fact that the most of the actuators, passive and active sensors are developed and manufactured in the People's Republic of China.

The authenticity of the devices and their metrological characteristics are questioned due to the general characteristics of the market for these devices. Certified hardware is not available in the price range offered by unauthorized online stores. The inhomogeneity of the gradation of the level of measurement accuracy to the proposed and publicly available datasheets is a big problem for modern entities.

Misleading the results of values from sensors to end users plays a significant role in information and technological security. An equally important role is played by the microprocessor programming image, code purity, and polling time. The irregulative role of relations associated with the ties of professional incompetence of the service personnel and the metrological negligence of the culture of production and service can even lead to tragic consequences. For example, a mismatch in the stability of the power supply of a gas sensor can not only lead to a sharp failure of it before a certain operating time for failure, but also lead not to correct decisions of the microprocessor system represented by a control device and a decision-making device.

However, the information entropy, expressed by the lack of clear standardization, and the lack of knowledge about the primary (operational) calibration and the conditions of use can be excluded by a number of proposed priori principles. First of all, it is necessary to implement the component that will be responsible for the analysis of input and output data, as well as for their processing. But first of all the contact bounce inherent in simple electronic keys and groups of switching devices shall be excluded. Single-ended devices in general have a weak stability under the influence of long-term operation [4].

In this connection, a proven digital circuitry solution such as the MS trigger is offered as an alternative or replacement to modern switches.

The use of a trigger is justified by the possibility of storing of the previous state and by the possibility of direct control (i.e. external interference). The rationale for such a possibility of controlling the device is associated with further considerations, the narrative of which is focused on the integration of an independent tool for analysing the process of reading information by sensors of the information-measuring environmental situation under consideration.

In addition, the use of triggers fits into the general model of relay-contactor systems, which are widely used due to their reliability in adverse environmental and climatic conditions [5].

An important advantage of these devices operating in binary logic is the ability to build triggers with any number of stable states, which simplifies the integration of multichannel output or switching devices. There are a lot of such devices in ecological systems. There are triggers executed based on
logical elements AND-NOT with 2, 3 and 4 stable states without loss of speed of operation and switching shown in figure 1.

![Figure 1](image1.png)

**Figure 1.** Demonstration of the scalability of the number of steady states based on trigger devices.

The MS trigger itself is a synchronous two-stage trigger with static control. In terms of circuitry, it is performed on two triggers: M (main) and S (auxiliary).

M-trigger perceives input information, and S-trigger captures the state of the trigger as a whole. In this case, both triggers can be of the same type (for example, RS- or D-triggers) or of different types. This property of versatility is very convenient at the stage of production of such devices.

The control connection between the M-trigger and S-trigger is carried out by two synchronous series, or by means of a restrictive inverter [6].

The functional diagram of a synchronous two-stage RS-trigger (see figure 2) with control performed as per MS-trigger, where a synchronous static RS trigger is used as an M-trigger and S-trigger is as follows:

![Figure 2](image2.png)

**Figure 2.** MS trigger with an inverter based on RS triggers based on the logical elements AND-NOT.

When the clock input value (which is important) is equal to CLK = "1" information is received in the M-trigger, but does not go to the S-trigger. If the value is reversed ("0"), information from the M-trigger is rewritten into the S-trigger.

This feature of the operation of the relay device can be used for the following purposes:
To analyse the previous output value of the temperature sensor/ humidity sensor/ gas analyser buffered in the S-trigger.

Buffering the output value for subsequent analysis (since triggers are devices with a frequent switching function, therefore, some often and strongly changing values of the sensor parameters (for example, with an instantaneous temperature change) may work in an erroneous way) [7].

Thus, the proposed circuit solution complements the relay-contactor component of the circuits of information-measuring systems. It will allow to diagnose equipment failures, based on the firm principle of maintaining the previous state of the output value more reliably (with a certain configuration).

3. Possibilities of use of computer algorithms for post-empirical reduction of measurement results

After receiving of a modified model of output values and determination of the system keys, the indispensable attributes of any information-analytical and metrological support of applied use shall be referred to. In particular to those attributes that are involved in the analysis of ecological systems.

The fact is not only in the variety of parameters but also in their general inhomogeneity. Something is measured in SI values, something is brought to the values already at the processing stage, and something (for example, voltage) has a four-vector physical element. All these is the concept of potential [8].

It have already been emphasized that in many ways, modern systems responsible for measuring parameters directly related to the biosphere, troposphere and hydrosphere of our world are far from ideal. However, this does not mean that they cannot be used.

It is practically impossible to influence the measurement result dealing with a final device designed in the form of a case that usually has its own memory.

However, it shall be noted that there is a uniform way of organizing of all these devices. If all these devices have memory, than it is shared one. There is a cyclic data reading program, and there is a data memory. Often, sensors have an external one. Its real analogue is a more powerful control system (see Fig. 3), which collects all the parameters and programs [9].

But why bring all the analysis of large data streams in a controlling system with shared memory (and most often the Smart Home and Ecosystem controllers have exactly the Harvad architecture), especially if it is about use of buffered states, which in general can already be issued visually, in the form of finite state machines and presented to the user. On the other hand, changing architecture to the opposite is very difficult, and it has its drawbacks. However, the principle of combining data memory and program memory that is convinient for analysis can be simulated as follows.

![Diagram of Von Neumann and Harvard Architectures](image-url)
Figure 3. A prior cycle of interaction of two architectures, applied to electronic computing technology.

Let's suppose that there is an experience of measuring the temperature of an object or a certain working cycle of the strain gauge. The output data has been sent through the data bus to one of the discrete inputs of the controller. Let's take into account that the clocking input of each key is connected to the output or reprogrammable (free) discrete outputs of the controller of the environmental measurement system.

Then, perhaps, the most logical variant of the classification of the controller's output data shall be applied even at the stage of generating a signal towards the controller. A logistic regression model shall be used [10].

Logistic regression is a classic machine learning component for solving a classification problem. ROC analysis is an apparatus for analysis of the quality of models. These methods are actively used to construct behaviour and response models in various dynamic environments. Logistic regression can be represented as a single-layer neural network with a sigmoidal activation function. The scalar values of this function are the logistic regression coefficients, and the polarization value is the constant of the regression equation.

What is it good for? First of all, it has an extremely wide map of use and rather simple computational complexity, in comparison with other machine learning algorithms. In fact, it is a special case of multiple regression, which has a target map in the analysis of the relationship between several independent variables and the dependent variable. This is exactly the case for almost all modern parametric measurement devices, except for sensors that operate on the "1-Wire" principle.

The method is based on the likelihood function, which is easy to calculate on the basis of at least several measurement iterations, for example, at the stage of initial commissioning, and which expresses the probability density of the joint occurrence of the sample results for a specific device and transmitted to the clocking input of each key and relay. In fact, this will mean that the system agrees to record a new previous state [11,12]. And since the regression algorithm is based on machine learning, then all data (as it is implied) will accumulate in the memory of the control system along with the parameters. Thereby these data emulate a kind of Neumann approach to architecture. At the final stage, everything is reduced to the fact that each new sensor value is simultaneously an instruction for action (expressing the imperative of the program memory) and the will of the data [13].

4. Findings
The use of a mixed - hardware (circuitry) and software-hardware (machine analysis and impact on clock inputs) approach to the implementation of relay-contactor logic, allows us to talk about an additional mechanism for monitoring and decision-making in relation to individual components of information and measurement systems.

Attention is paid to the circumstances that require a more responsive and precise instrument for the continuous analysis of the stability of the operation of devices, their electrical characteristics. The possibility of transition to a more advanced and already studied method of switching functionality is demonstrated, the applied value of buffering key states of sensors for further dispatching and monitoring in rapidly changing environments is emphasized.

References
[1] Azen R and Walker C M 2011 Categorical Data Analysis for the Behavioral and Social Sciences (Biostatistical Methods: The Assessment of Relative Risks) 2nd ed. Wiley
[2] Aleksenko A G 1971 Microcircuit technique fundamentals (Moscow: Soviet radio publisher) 350 p
[3] Bernard Cohen and Howard Aiken 2000 Portrait of a computer pioneer (Massachusetts: The MIT Press) 53 p
[4] Buslenko N P and Shreider Yu A 1961 Statistical test method (Monte Carlo) and its
implementation on digital computers (Moscow: FIZMATLIT publisher) 226 p
[5] Izrael Yu A 1979 Ecology and environmental state control (Leningrad: Gidrometeoizdat publisher) 376 p
[6] Izrael Yu A 1974 Global Monitoring System. Forecast and assessment of environment. Fundamental principles of monitoring Meteorologiya i gidrologiya publisher 7 pp 3-8
[7] Metlev V D, Panchenko L A, Riznichenko G Yu and Terekhin A T 2009 Theory of Probability and Mathematical Statistics. Mathematical models (Moscow: Akademiya) 320 p
[8] Pospelov D A 1974 Logical methods of analysis and synthesis of schematics 3rd edition (Moscow: Energiya publisher) 368 p
[9] Hilbe J M 2009 Logistic Regression Models Chapman&Hall/CRC Press
[10] Hosmer D 2013 Applied logistic regression. Hoboken, New Jersey: Wiley
[11] Yumashev A, Ślusarczyk B, Kondrashev S and Mikhaylov A 2020 Global indicators of sustainable development: Evaluation of the influence of the human development index on consumption and quality of energy Energies 13(11) 2768 doi: 10.3390/en13112768 https://www.mdpi.com/1996-1073/13/11/2768
[12] Zakharchenko N V, Hasanov S L, Yumashev A V, Admakin O I, Lintser S A and Antipina M I 2018 Legal rationale of biodiversity regulation as a basis of stable ecological policy Journal of Environmental Management and Tourism 9(3) pp 510-23
[13] Vasilev V L, Gapsalamov A R, Akhmetshin E M, Bochkareva T N, Yumashev A V and Anisimova T I 2020 Digitalization peculiarities of organizations: A case study Entrep 7 3173-90