

PAPER • OPEN ACCESS

## Industrial Informatics: Towards the Researches and Modelling of Structural Materials' Properties

To cite this article: E E Kovshov *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **476** 012016

View the [article online](#) for updates and enhancements.

# Industrial Informatics: Towards the Researches and Modelling of Structural Materials' Properties

**E E Kovshov<sup>\*</sup>, A A Kosach and V S Popov**

JSC «NIKIMT-Atomstroy», Moscow, Russian Federation

<sup>\*</sup>E-mail: eugene.kovshov@gmail.com

**Abstract.** Considers the possibility of applying the principles of Industrial Informatics to optimize the study and modelling of various structural materials and products on the example of intelligent software and hardware platform based on mathematical apparatus and modules. The ways of practical implementation of mathematical and bio-inspired algorithms and models for the developed software applications in the corporate information system of industrial enterprise paradigm-based Industry 4.0 are proposed.

## 1. Introduction

Industrial Informatics as an independent scientific and engineering direction appeared at the junction and as a result of the rapid development of science, technology and computer science. The term "industrial" refers to the creation of real industrial software applications, and the term "informatics" refers to the infrastructure that provides development, implementation and maintenance throughout the life cycle of existence of it. In addition, Informatics as a science offers tools and methods for analysis, processing, transformation and transfer of information. Industrial Informatics focuses primarily on the automation based on knowledge as a means of re-engineering and modification of production processes in industry.

Industrial Informatics is not limited to metal working, for example, there are such branches of knowledge as computer control systems, robotics, intelligent video surveillance and image processing systems, as well as data collection of different nature and processing of multidimensional signals, where mathematical methods and tools of Informatics are widely used.

Industrial Informatics has a set of techniques and practices that use information analysis, processing and dissemination to achieve more meaningful results in terms of efficiency, reliability and/or safety in an industrial environment. The field of industrial Informatics has become one of the key areas for intelligent control and digital production technologies.

Information Technology (IT) design tools for different industries vary widely depending on the nature of the industry and, accordingly, the business processes being implemented.

For example, the use of IT in manufacturing industry includes modeling business processes, planning and production control, distributed information system inventory management and knowledge management for applied research.

Industrial Informatics – "building", built on the foundation of "classical" Informatics using the theory of algorithms, fuzzy logic, artificial neural networks, artificial immune systems, evolutionary and genetic algorithms, industrial design of information systems, as well as accumulated considerable practical experience in the field of theoretical and applied Informatics. Industrial design processes are



defined as a set of logically related tasks performed to achieve a specific result, with the industrial production process defined based on the customer's requirements and business interests.

## 2. Digital intelligent platform

One of the main trends in the development of application software in the structure of digital production was the transition from monolithic construction of information systems to modular, which greatly facilitates the scaling of the software.

The last stage in the development of modular programming was the service-oriented architecture (SOA) of information systems [1]. SOA is a modular approach to software development based on the use of distributed, loosely coupled standardized components (services). Services are compatible, flexible and reusable. Accordingly, all the functionality of the system is implemented in components that are built in such a way that they are nearly unrelated to each other. The introduction of information systems based on SOA becomes preferable with the growing scale of data and calculations.

This intelligent software platform for data collection and processing with non-destructive testing is based on service-oriented architecture of software solutions and is one of the information subsystems in the structure of integrated digital production.

Developed and tested (as part of the SOA platform) intelligent software platform for data collection and processing of the automated information-measuring system of leakage testing [2]. The services perform various application tasks (e.g. image recognition, strength test calculation) and are combined with each other only to solve the problems identified by their application shown on the model (figure 1).

It should be noted that business processes serve as a mechanism for ensuring the interaction of services with each other. A business process is a set of interrelated tasks and it controls the flow of events, calls and coordinates services, and creates a context for their interaction. A business process is an abstract mechanism that does not depend on how services are implemented and what logic they carry. The main task of the business process is the organization of services for its effective implementation.

The orchestration module is designed to select a business process that interacts with different services. It defines a mechanism for service interaction based on business process logic and represents a sequence of actions performed by services. The enterprise bus (in this case, ESB) provides an environment for interacting with the service.

One of the most efficient ways to implement a service-oriented approach is to use web-services, where the service-oriented model covers all levels of data management in a software platform based on SOA. It also has no distribution to suppliers and consumers, as all the necessary information about the services is in the register. For more complex and in-depth data analysis, services located in a public, private, or hybrid cloud can be used.

Open ESB corporate service bus was used for implementation of information-measuring control system [3]. Services are implemented in the form of web-services, the functional part of which is written in Java. The system elements exchange messages based on the SOAP Protocol [4]. To store data and calculation results an open-source relational database management system is used. The use of these technologies allows you to analyze and filter data, including in the main Microsoft Office software (for example, Microsoft Excel, Microsoft Access).

For the exchange of information between parts of the information-measuring system applied OPC servers (OPC UA, OPC DA) and the communication protocols Modbus RTU and Modbus TCP. OPC is a technology based on the use of a single device management interface [5], this is especially important for automating remote control systems. Often, developers of automated systems have to include a lot of ready-made drivers in the control system or use the tools to develop original communication protocols with non-standard devices of the lower level. And when they replace a device with a similar one from another manufacturer, they need to significantly change the software to

control that device. A single OPC interface allows different software modules produced by different companies to interact with each other.

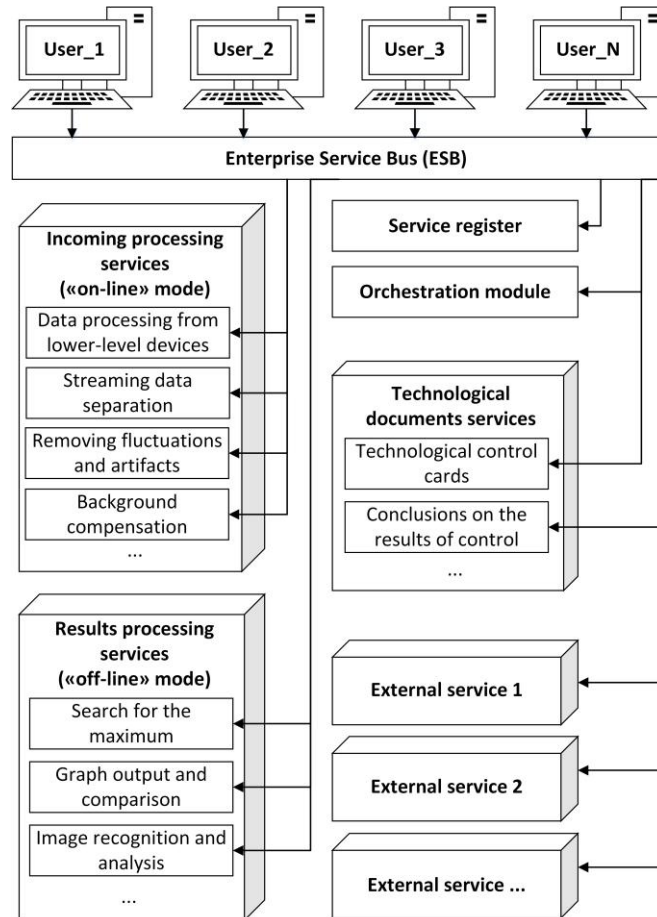


**Figure 1.** The layout of the software platform based on SOA.

### 3. Mathematical apparatus

This software platform from our point of view is applicable for the selection of structural materials with certain properties in the design of products and to reduce the number of options to be modeled and calculated. Modeling and calculation of products requires significant time and computer resources, and the preliminary selection of the most preferred options from the set reduces these costs. It is also possible to significantly improve certain characteristics of the product, making it, for example, from composite materials or layers of different materials, and the use of algorithms for optimization allows you to select solutions from a very wide range of possible.

At the moment, these tasks are performed mainly using genetic algorithms [6]. Genetic algorithm is an adaptive method of finding solutions. It uses the mechanisms of genetic inheritance, natural selection, biological terminology and basic concepts of linear algebra. The main idea of the genetic algorithm is the organization of natural selection among a variety of solutions to find the extremum of the objective function [7].

Besides this method, you can solve the problem of selecting the required structural material using a neural network or cluster analysis. These methods also allow you to select the most appropriate option from the set with a large number of specified conditions.

Depending on the task, it is advisable to optimize the selection of material on the basis of a neural network module with training with a teacher. The neural network consists of a multilayer perceptron [8], which includes:  $n$  layers of  $k$  neurons transmitting information through channels with weights ( $w_{ij}$  - the weight of the connection between neurons  $i$  and  $j$ ),  $d$  inputs and one output, which gives the result (figure 2).

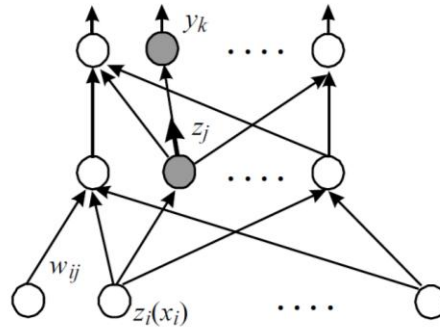


**Figure 2.** A fragment of the artificial neural network.

When information is received at the inputs, it is processed by the first layer of neurons, each of which calculates the weighted sum of its inputs:

$$a_i = \sum_n w_{ij} z_i \quad (1)$$

where  $z_i$  is the output of neuron  $i$  and the input of neuron  $j$ , then the output of neuron  $j$  is the conversion of  $a_j$  by the threshold function of this neuron:

$$z_j = g(a_j) \quad (2)$$

The training of an artificial neural network is to determine the weights for each bonds. When training with the teacher, a sample of input vectors  $\{x_n\}$  is used, for which the desired result of the neural network  $t \in \{t_n\}$  is known (for example, the results of modeling products from certain materials), which allows you to effectively adjust the network weights.

Also, for the selection of materials with the necessary properties, SVM-classifier (SVM - Support Vector Machines) is used [9], which implements training on precedents (analog of training with a teacher). This classifier is advisable to use for the distribution of a large number of solutions (in this case, with respect to materials for the manufacture of certain products) for classes that meet a set of specified characteristics.

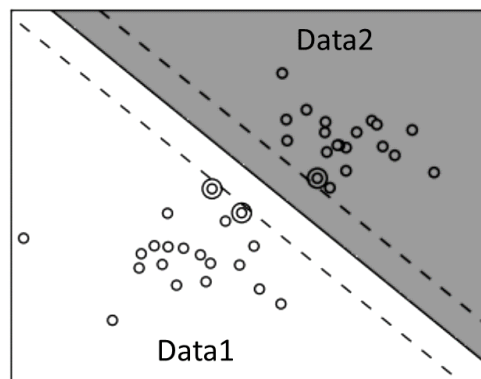


**Figure 3.** The result of separating two-dimensional data by SVM classifier.

Consider the two-dimensional option. Each characteristic of classified objects is represented as a  $q$ -dimensional vector of numerical values of characteristics  $z_i = (z_{i1}, z_{i2}, \dots, z_{iq})$ , where  $z_{iq}$  is the numerical value of characteristic  $q$  for object  $i$  and each of these vectors must belong to only one of the two classes. Then the search for the separating hyperplane with the maximum gap in this space is carried out (figure 3).

There is a data set  $\{(z_1, y_1), \dots, (z_n, y_n)\}$ , where each object  $z_i$  of the set  $Z$  corresponds to the number  $y_i$  of the set  $Y$ , which takes the value "-1" or "+1", depending on the class to which the object  $z_i$  belongs. Required to create a classifier  $F$  that associates a random object from the set  $Z$  with a class from the set  $Y$  using the kernel function  $k$ . As a kernel function that allows you to separate objects of different classes, different functions are used (for example, linear or polynomials).

After training, the classification function is determined:

$$f(z) = \sum_{i=1}^N \alpha_i \cdot y_i \cdot \kappa(z_i, z) + b \quad (3)$$

where  $k$  is a function of the classifier kernel;  $b$  is a parameter that determines the shortest distance from the beginning to the hyperplane;  $\alpha_i$  - Lagrange multiplier;  $y_i$  is a classification solution.

When developing an SVM classifier, the source data (training set) is randomly divided into training and test samples. The training sample is used for training the classifier. The test sample (some quantity of objects of the training set) is not used for training the classifier, and is used to verify the accuracy of the classifier. The result of learning the SVM algorithm is a classifying function.

The decision of classification, which associates the object with the class with a "-1" or "+1", is determined by the formula:

$$F(z) = \text{sign}(f(z)) = \text{sign}\left(\sum_{i=1}^N \alpha_i \cdot y_i \cdot \kappa(z_i, z) + b\right) \quad (4)$$

Denote for a training set  $U$  containing a set of  $Z$  objects the classification rule  $A(z)$ . Then the following values the number of outcomes of the correct classification can be determined:

- number of correctly defined objects of class "+1":

$$TP(A, Z^*) = \sum_{z=Z^*} (A(z) = +1 \wedge y^*(z) = +1) \quad (5)$$

- number of correctly defined objects of class "-1":

$$TN(A, Z^*) = \sum_{z=Z^*} (A(z) = -1 \wedge y^*(z) = -1) \quad (6)$$

And the values that determine the number of incorrect classification outcomes [9]:

- number of incorrectly defined objects of class "+1":

$$FP(A, Z^*) = \sum_{z=Z^*} (A(z) = +1 \wedge y^*(z) = -1) \quad (7)$$

- number of incorrectly defined objects of class "-1":

$$FN(A, Z^*) = \sum_{z=Z^*} (A(z) = -1 \wedge y^*(z) = +1) \quad (8)$$

Relative performance indicators of the classifier can also be formulated for further calculation of the classification quality:

- sensitivity (completeness) is the ratio of correctly defined objects of class "+1" to objects really belonging to class "+1":

$$R(A, Z^*) = \frac{TP(A, Z^*)}{TP(A, Z^*) + FN(A, Z^*)} \quad (9)$$

- accuracy is the ratio of correctly defined objects of class "+1" to objects assigned to class "+1" in the classification process:

$$P(A, Z^*) = \frac{TP(A, Z^*)}{TP(A, Z^*) + FP(A, Z^*)} \quad (10)$$

To assess the quality of the classification, taking into account the accuracy and completeness, it is advisable to use combined metrics. For example, the  $F_\beta$ -measure:

$$F_\beta(A, Z^*) = \frac{(\beta^2 + 1)P(A, Z^*) \cdot R(A, Z^*)}{\beta^2 P(A, Z^*) + R(A, Z^*)} \quad (11)$$

where  $\beta$  is the coefficient reflecting the expected value of accuracy relative to completeness.

Or balanced metric ( $F_1$ ):

$$F_1(A, Z^*) = \frac{2P(A, Z^*) \cdot R(A, Z^*)}{P(A, Z^*) + R(A, Z^*)} \quad (12)$$

The balanced metric is a special case of the  $F_\beta$ -measure ( $\beta=1$ ) and is equal to the mean harmonic between the accuracy and completeness indices.

The value of  $F_1$  tends to 1 if the accuracy and completeness are close to 1, which shows good quality of the classification. Thus, knowing these indicators it is possible to establish in the course of training whether the given classifier for a certain task is correctly picked up and whether in sufficient volume training was carried out.

#### 4. Conclusion

The current stage of development of society is characterized by a significant impact on it of information and telecommunication technologies, which are increasingly and deeply seep into all spheres of human activity without exception, ensure the dissemination of information in society and, at the same time, forming a global information space. An integral and important part of these processes is the penetration of modern information and communication technologies in the field of industrial production and digital materials science as well.

Significant simplification and cheapening of the industrial production brings widespread use of information technology in combination with various means of mathematical analysis and adaptive methods of finding solutions, including in the modeling and study of the structural materials' properties.

#### References

- [1] Vasiliev Y 2007 *SOA and WS-BPEL* (Packt Publishing Ltd.32)
- [2] Kosach A and Kovshov E 2017 Development of a computer based problem-oriented automated remote leakage detection control system *ITM Web of Conference* **10** 01005
- [3] Sun GlassFish Enterprise Service Bus <http://download.oracle.com/docs/cd/E19182-01/index.html>
- [4] Kosach A and Kovshov E 2017 Automated remote computer monitoring of environmentally hazardous products *The 14th International Conference «Application of Contemporary Non-Destructive Testing in Engineering»* (Slovenia: University of Ljubljana) pp 95-102
- [5] Mahnke W, Leitner S and Damm M 2009 *OPC Unified Architecture* (Springer)
- [6] Leung M and Nevill G 1993 Genetic Algorithms for Preliminary 2-D Structural Design *AIAA Journal* **94** 2287-2291
- [7] Punch W, Averill R and Goodman E 1995 Design using Genetic Algorithms - Some Results for Composite Material Structures *IEEE Expert – Intelligent Systems and their Applications* pp 42-49
- [8] Kandel E, Schwartz J and Jessell T 2000 *Principles of neural science* (Appleton & Lange)
- [9] Demidova L, Sokolova Yu and Nikulchev E 2015 Use Of Fuzzy Clustering Algorithms' Ensemble For SVM Classifier Development *International Review on Modelling and Simulations* **8** 446-457