Method Detection Audit Data Anomalies on Basis Restricted Cauchy Machine

Tetiana Neskorodieva[1], Eugene Fedorov[1], Oleksii Smirnov[2], Kostiantyn Rudakov[3], and Anastasiia Neskorodieva[1]

[1] *Vasyl' Stus Donetsk National University, 21 600-Richcha str., 21021, Vinnytsia, Ukraine*
[2] *Central Ukrainian National Technical University, 8 Universytetskyi ave., 25006, Kropyvnytskyi, Ukraine*
[3] *Cherkasy State Technological University, 460 Shevchenko ave., 18006, Cherkasy, Ukraine*

Abstract

The paper presents a method for the anomalies detection in waste-free production audit data based on the neural network model of Gauss-Bernoulli bidirectional restricted Cauchy machine (BRCM). The purpose of the work is to increase the efficiency of audit data analysis of waste-free production on the basis of the neural network model of anomalies detection without the use of the marked data that simplifies audit. To achieve this goal, the following tasks have been set and solved: offered model of generalized multiple transformations of audit data in the form of a two-layer neural network. Cauchy offered neural network model of Gauss-Bernoulli bidirectional restricted Cauchy machine possesses a heteroassociative memory; works real data; has no restrictions for memory capacity; provide high accuracy of anomalies detection; uses Cauchy's distribution that increases the speed of convergence of a method of parametrical identification. To increase the speed of Gauss-Bernoulli parametric identification of a bidirectional restricted Cauchy machine, a parametric identification method was developed to be implemented on a GPU using CUDA technology. The offered method allows increasing training speed by approximately proportional to the product of numbers of neurons in the hidden layer and power of a training set. The made experiments confirmed the operability of the developed software and allow to recommend it for use in practice in a subsystem of the automated analysis of DSS of audit for detection of anomalies.

Keywords

Audit, mapping by neural network, Gauss-Bernoulli bidirectional restricted Cauchy machine, anomalies detection.

1. Introduction

Nowadays the scientific and technical issue of the modern information technologies in financial and economic sphere is creation methodology forming of the decision support systems (DSS) at the enterprises audit in the conditions of IT application on enterprises and with the use of information technologies. Modern automated DSS audit are based on the automated analysis of the large volumes of data about financial and economic activity and states of enterprises with the multilevel hierarchical structure of heterogeneous, multivariable, multifunction connections, intercommunications and cooperation of objects of audit. The tasks automated DSS audit are expansion of functional possibilities, increase of efficiency and universality of IT-audit [1].

CPITS-II-2021: Cybersecurity Providing in Information and Telecommunication Systems, October 26, 2021, Kyiv, Ukraine
EMAIL: t.neskorodieva@donnu.edu.ua (T. Neskorodieva); y.fedorov@chdtu.edu.ua (E. Fedorov); dr.smirnovoa@gmail.com (O. Smirnov); k.rudakov@chdtu.edu.ua (K. Rudakov); t.neskorodieva@donnu.edu.ua (A. Neskorodieva)
ORCID: 0000-0003-2474-7697 (T. Neskorodieva); 0000-0003-3841-7373 (E. Fedorov); 0000-0001-9543-874X (O. Smirnov); 0000-0003-0000-6077 (K. Rudakov); 0000-0002-8591-085X (A. Neskorodieva)

2. Problem Statement

Let for model of detection of anomalies the training set be set $S = \{(\mathbf{x}_m^{in}, \mathbf{x}_m^{out}, \mathbf{d}_m^{in}, \mathbf{d}_m^{out})\}$, $m \in \overline{1, M}$, where $\mathbf{x}_m^{in}$ is m - y raw materials vector, $\mathbf{x}_m^{out}$ is m - y vector of finished goods, $\mathbf{d}_m^{in}$ is m - y the expected reference vector of raw materials, $\mathbf{d}_m^{out}$ is m - y the expected reference vector of finished goods.

Then a problem of increase in accuracy of detection of anomalies on Gauss-Bernoulli's model of the bidirectional limited machine of BRCM $g(\mathbf{x}^{in}, \mathbf{x}^{out}, \mathbf{w})$, where $\mathbf{x}^{in}$ is raw materials vector, $\mathbf{x}^{out}$ is the vector of finished goods, w is a vector of parameters, is represented as a stay problem for this model of such vector of parameters $\mathbf{w}^*$, which meets criterion $F = \frac{1}{M}\sum_{m=1}^{M}(g(\mathbf{x}_m^{in}, \mathbf{x}_m^{out}, \mathbf{w}^*) - (\mathbf{d}_m^{in}, \mathbf{d}_m^{out}))^2 \to \min$.

3. Literature Review

Currently, the analytical procedures used during the audit are based on data mining techniques [2, 3]. Automated DSS audit means the automatic forming of recommendable decisions, based on the results of the automated analysis of data, that improves quality process of audit. Unlike the traditional approach, computer technologies of analysis of data in the system of audit accelerate and promote the process accuracy of audit, that extremely critical in the conditions of plenty of associate tasks on lower and middle levels, and amounts of indexes and supervisions in every task.

The development of methods of estimation and prediction [4, 5], formation of generalized associative relationships [6] are described in the works of the authors of this article. The goals of creating these methods: reducing the computational complexity for simple tasks (a single mapping of elements or sub-elements of the audit subject area), automatic structural identification, increasing the accuracy for com-plex tasks (compositions of mappings of elements or sub-elements of the audit sub-ject area) and the possibility of applying these methods for the generalized analysis of elements and sub-elements of the audit subject area (Table 1).

The choice of model in the audit DSS depends on:

1. Characteristics of the audit data type (time series data, spatial data as mappings).
2. Audit level (upper middle, lower).
3. Audit tasks (internal, external).
4. The type of analysis tasks (detection of anomalies, structural analysis, assessment of indicators).
5. The characteristics of the enterprise (large, medium, small) and the type of activity (industry) at the top level.
6. Characteristics of sets and subsets of operations at lower levels (numerological, quantitative, semantic, logical).

This choice is schematically formalized in the form of a binary decision tree for choosing a neural network data audit model (see Fig. 1).

The proposed logical-neural network method makes it possible to automate the process of data analysis in the audit DSS and optimize it depending on the characteristics of the audit process and the audit object. One of the main tasks of data analysis of the audit subject area is the identification of anomalies. Let's consider the existing types of anomalies and methods of their operation.

Types of anomalies [7–9]:

- Point (are provided by points in character space).
- Contextual (usually a point of a time series or the rarefied data which depends on the environment).
- Collective (the section of a time series or the rarefied data).

Methods of detection of anomalies [7–9]:

1. Approach on the basis of rules (logical approach):
    I. methods on the basis of associative rules with classification and without classification (for example, the Apriori method);

II. methods on the basis of a decision tree with classification (for example, a method of the isolated wood).

Table 1

Comparative analysis of intelligent analysis methods in audit tasks

The economic content of the display	Model of processing elements of the subject area, Features of the model or method	Purpose of processing elements of the subject area	Advantages disadvantages of the model or method
Payment - delivery of raw materials	Modified Liquid State Machine, one- dimensional hidden layer, parameter identification based on matrix pseudoreversion [1]	Evaluation and prediction of indicators of raw material supplies (by type) based on the values of payment indicators in a direct check of the display	Reducing computational complexity, improving the forecast accuracy
Settlements with suppliers-customer settlements	A neural network model based on a gateway recurrent unit. For parametric identification of this model, adaptive cross entropy (a combination of random and directional search) is faster to learn but less accurate than in [1] because the pseudoreversal is not paralleled	Evaluation of indicators of settlements with customers on the basis of values of indicators of settlements with suppliers in a direct verification of mapping	Reducing computational complexity, improving the forecast accuracy
Settlements with suppliers - settlements with customers (a composition of mappings between a set of input and output data)	Forward-only counterpropagating neural network, which is a nonrecurrent static two-layer ANN [2], assumed that the audit indicators are noisy with Gaussian noise	Construction of generalized associative relationships for generalized analysis tasks (in the forward direction)	Automating the formation of generalized features of audit sets and their mapping by means of a forward-only counterpropagating neural network the number of pairs (neurons in the hidden layer N1) is set manually
Release of raw materials - posting of finished products (a composition of mappings between a set of input and output data)	Bidirectional counterpropagating neural network, which is a nonrecurrent static two-layer ANN BCPNN	Construction of generalized associative relationships for generalized analysis tasks (in the forward and backward direction)	Automating the formation of generalized features of audit sets and their mapping by means of a bidirectional counterpropagating neural network the number of pairs (neurons in the hidden layer) is set manually
Payment - delivery of raw materials	Modified Liquid State Machine, one- dimensional hidden layer, parameter identification based on matrix pseudoreversion [1]	Evaluation and prediction of indicators of raw material supplies (by type) based on the values of payment indicators in a direct check of the display	Reducing computational complexity, improving the forecast accuracy

2. Approach on the basis of ANN:

   I. ANN without classification (for example, the one-class SVM (support vector machine), ANN associative memory (for example, the autoencoder, SOFM (self-organizing feature map), Hopfield neural network, Boltzmann machine), ANN of the forecast of a time series (for example, NARNN (non-linear autoregressive neural network), NARMANN (nonlinear autoregressivemoving average neural network), SRN (simple recurrent network), BRNN (bidirectional recurrent neural network), LSTM (long short-term memory), BiLSTM, GRU (gated recurrent unit), BiGRU));

II. ANN with classification (for example, MLP (multilayer perceptron), RBFNN (radial-basis function neural network)).

3. Approach on the basis of Bayes's networks with classification
4. Approach on the basis of a clustering:
   I. clustering on the basis of centroid (for example, a method of k-means) or distributions (for example, the EM (expectation–maximization) method);
   II. clustering on the basis of medoid (for example, the PAM (partitioning around medoids) methods, a subtractive clustering);
   III. density clustering (for example, DBSCAN (density-based spatial clustering of applications with noise) methods, OPTICS (ordering points to identify the clustering structure methods).

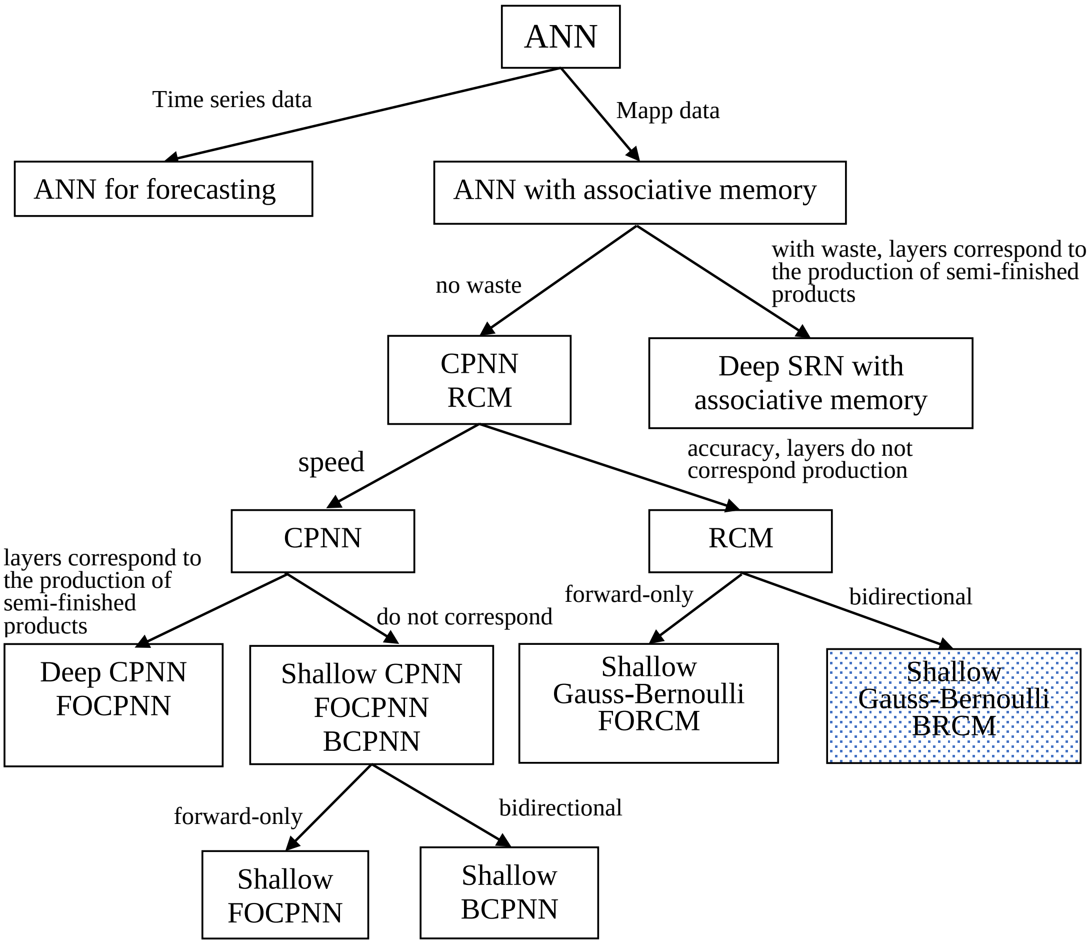

Figure 1: Binary decision tree of neural network model selection for data analysis.

5. Approach on the basis of the neighborhood (metric approach) (for example, methods of the k-nearest neighbors, LOF (local outlier factor))
6. Approaches on the basis of distributions:
   a. Parametrical approach on a basis:
      I. Gaussian distributions (for example, MCD (minimum covariance determinant) method);
      II. mixtures of distributions (for example, HMM (hidden Markov models), GMM (Gaussian mixture models)).
   b. Nonparametric approach on a basis:
      I. histograms;
      II. functions of a kernel (for example, Parzen window method).
7. Approach on the basis of regression model (for example, the Box-Jenkins meth-od)

8. Approach on the basis of the spectral theory (matrix decomposition) (for example, PCA (principal component analysis) method)
9. Approach on the basis of information theory (entropy).

Now the most popular is approach of detection of anomalies based on neural networks.

Disadvantages of the one-class SVM is restriction for quantity of support vectors. A disadvantage of ANN of the forecast of a time series is that they require existence of a time series. A disadvantage of ANN with classification is the requirement to classify anomalies that is not always possible owing to labor input of obtaining the marked data on each type of anomalies. Therefore, in our work, we chose ANN with associative memory.

Traditional neural networks with an associative memory are:

1. Neural networks only with heteroassociative memory (for example, FOCPNN (forward-only counter propagation neural network) [10], PCANN (principal com-ponent analysis neural network) [11], ICANN (independent component analysis neural network) [12], CMAC (cerebellar model articulation controller) [13].
2. Neural networks only with an autoassociative memory (for example, the autoencoder [14], SBN (sigmoid belief network) [15], Helmholtz machine [16], SOFM [17], LVQNN (learning vector quantization neural network) [18], RCAM (recurrent correlation associative memory) [19], Hopfield neural network [20], Gauss machine [21], BSB (brain-state-in-box) [22], Hamming neural network [23], ART (Adaptive resonance theory) [24].
3. ANN with a heteroassociative and autoassociative memory (for example, BCPNN (bidirectional counter propagation neural network) [25], BAM (bidirectional associative memory) [26], Boltzmann machine [27]).

The majority of neural networks with an associative memory possess some or more shortcomings:

1. Do not possess at the same time autoassociative and heteroassociative memory,
2. Do not work with material data.
3. Have no high capacity of an associative memory.
4. Have no high accuracy.
5. Have high computing complexity.

In this regard, creation of a neural network which will allow to eliminate the specified defects is relevant.

The purpose of work is increase in efficiency of data analysis of audit of waste-free production on the basis of neural network model of detection of anomalies with-out use of the marked data that simplifies audit [28–30].

For achievement of the goal, it is necessary to solve the following problems:

- Offer neural network model of detection of anomalies.
- Select criterion for evaluation of efficiency of neural network model of detection of anomalies.
- Offer a method of parametrical identification of neural network model of detection of anomalies.
- Execute numerical researches.

4. Block Diagram of Neural Network Model of Detection of Anomalies

In this paper, the structure of the data transformation model is determined based on the production structure. It is assumed that the transformation of raw materials into finished products in one step without waste without intermediate products. Each type of raw material is used in the production of one or more types of finished products. The production structure for each planning period (month, quarter, year) is deter-mined on the basis of long-term contracts and short-term (in particular urgent) orders. The production plan is decomposed into quantization periods of the planning period, taking into account the production capacity for different types of products.

In this case, the transformation of these raw materials into finished products for the planning period can be represented in the form of a two-layer neural network. The number of neurons in the input layer is equal to the number of raw materials used in production. The number of neurons in the output layer is equal to the number of types of finished products. The input values are the amount of raw materials by type, the output of the network is the finished product values for the planning period or the quantization period.

When taking into account the supply of raw materials and the release of finished products for the quantization period, the amount of recorded released raw materials and released finished products is made up of the actual values of indicators and the difference in residuals at the beginning of the quantization period. These balances are not calculated at the end (beginning) of each quantization period and not reflected in accounting systems, balances are determined only at the beginning and end of the planning period).

The detecting anomalies problem during the release (write-off) of raw materials for the production of finished products for the periods of quantization of the verification period (in particular, the periods of quantization can be chosen equal to the production cycle - shift, day, several days) is formulated as follows. Determine the quantization periods for which the structure of consumption of raw materials is significant (the level of materiality is set by the decision maker) is higher or lower than the average values for the period according to the data of the release of raw materials into pro-duction and the posting of products.

This data transformation model is used to create an anomaly detection model that can be represented as a two-layer neural network. The number of neurons in the visible layer is equal to the sum of the number of raw materials used in production and the number of types of finished products. Thus, the number of raw materials and finished products is supplied to the visible layer by type for the planning period or the quantization period of the verification period. To train the neural network, the "correct" data are used (the formation of which has been verified). Data that are subject to verification are used as control data.

The block diagram of model of Gauss-Bernoulli BRCM (bidirectional restricted Cauchy machine) [6] which is recurrent ANN and contains one visible layer and one hidden layer (Fig. 2).

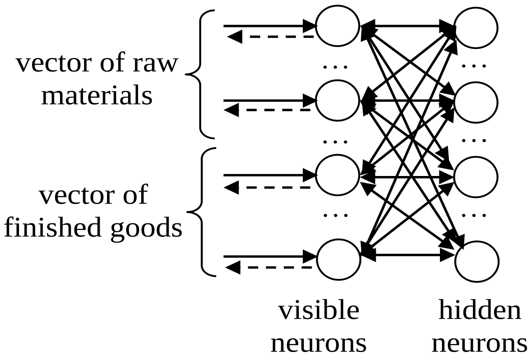

Figure 2: Block diagram of model of Gauss-Bernoulli BRCM (bidirectional restricted Cauchy machine)

Gauss-Bernoulli's components of BRCM are:

- Stochastic visible neurons which state is described based on Gaussian distribution in a form

$$x_j = \mu_j + \sigma_j N(0,1),$$

where μ_j is mathematical expectation, which characterizes the average value of indicators of supplied raw materials or capitalized products, σ_j is a mean square deviation (if the training a vector are normalized and centered, then σ_j=1), which characterizes the variance of the difference in residuals at the beginning of the quantization period, $N(0,1)$ is the function returning standard normally distributed random number.

Transition probability j-th a stochastic neuron in a state α is defined in a form

$$P_j = \frac{1}{\sigma_j\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{\alpha-\mu_j}{\sigma_j}\right)\right).$$

- The stochastic hidden neurons which state is described on the basis of Bernoulli's distribution in a form

$$x_j = \begin{cases} 1, & \text{with probability } P_j, \\ 0, & \text{with probability } 1 - P_j. \end{cases}$$

Transition probability j-th a stochastic neuron in a state 1 is defined in a form

$$P_j = \frac{1}{2} + \frac{1}{\pi}\arctan\left(\Delta E_j\right),$$

where ΔE_j is increment of energy of ANN at state change j-th a stochastic neuron with 0 on 1.

Gauss-Bernoulli's BRCM advantages:

1. Unlike the majority ANN possesses at the same time autoassociative and heteroassociative memory.
2. Unlike a bidirectional associative memory and Boltzmann machine works real data.
3. Unlike a bidirectional associative memory and Boltzmann machine has no restrictions for memory capacity.
4. Unlike a bidirectional associative memory provide big accuracy.
5. Unlike Boltzmann machine has smaller computing complexity.

5. Neural Network Model of Detection of Anomalies

Positive phase (Step 1–3).

1. Initialization of a state of the visible neurons corresponding to raw materials $\mathbf{x}1^{in} = \mathbf{x}^{in}$
2. Initialization of a state of the visible neurons corresponding to finished goods $\mathbf{x}1^{out} = \mathbf{x}^{out}$
3. Calculation of a state of the hidden neurons $\left(j \in \overline{1, N^h}\right)$.

$$P_j = \frac{1}{2} + \frac{1}{\pi}\arctan\left(b_j^h + \sum_{i=1}^{N^{in}} w_{ij}^{in-h}\frac{x1_i^{in}}{\sigma_i^{in}} + \sum_{i=1}^{N^{out}} w_{ij}^{out-h}\frac{x1_i^{out}}{\sigma_i^{out}}\right).$$

$$x1_j^h = \begin{cases} 1, & P_j \ge U(0,1), \\ 0, & P_j < U(0,1). \end{cases}$$

where $U(0,1)$ is the function returning uniform distributed random number in the range $[0,1]$.

Negative phase (Steps 4 and 5).

4. Calculation of a state of the visible neurons corresponding to raw materials $\left(j \in \overline{1, N^{in}}\right)$

$$\mu_j^{in} = b_j^{in} + \sigma_j^{in}\sum_{i=1}^{N^h} w_{ij}^{in-h} x1_i^h,$$

$$x2_j^{in} = \mu_j^{in} + \sigma_j^{in} N(0,1).$$

5. Calculation of a state of the visible neurons corresponding to finished goods $\left(j \in \overline{1, N^{out}}\right)$

$$\mu_j^{out} = b_j^{out} + \sigma_j^{out}\sum_{i=1}^{N^h} w_{ij}^{out-h} x1_i^h,$$

$$x2_j^{out} = \mu_j^{out} + \sigma_j^{out} N(0,1),$$

where b_j^h is bias for j-th of a neuron of the hidden layer, b_j^{in} is bias for j-th of a neuron of the visible layer corresponding to raw materials, b_j^{out} is bias for j-th of a neuron of the visible layer corresponding to finished goods, w_{ij}^{in-h} is connection weight from the neuron i-th in a visible layer corresponding to raw materials to j-th to a neuron of the hidden layer, w_{ij}^{out-h} is connection weight from the neuron i-th in a visible layer corresponding to finished goods to j-th to a neuron of the hidden layer, N^h is number of neurons in the hidden layer, N^{in} is the number of the neurons in a visible layer corresponding to raw materials, N^{out} is the number of the neurons in a visible layer corresponding to finished goods.

6. Choice of Criterion for Evaluation of Efficiency of Neural Network Model of Detection of Anomalies

In work for training of the BRCM model the function of the purpose which means the choice of such values of a vector of parameters is selected $\mathbf{w} = (w_{11}^{in-h}, ..., w_{N^{in}N^h}^{in-h}, w_{11}^{out-h}, ..., w_{N^{out}N^h}^{out-h})$, which deliver a minimum of a root mean square error (the differences of a sample on model and a test sample)

$$F = \frac{1}{M(N^{in} + N^{out})} \sum_{m=1}^{M} \left(\left\| \mathbf{x}2_m^{in} - \mathbf{d}_m^{in} \right\|^2 + \left\| \mathbf{x}2_m^{out} - \mathbf{d}_m^{out} \right\|^2 \right) \to \min_{\mathbf{w}},$$

where $\mathbf{x}2_m^{in}$ is m-th an evaluation vector of raw materials on model, $\mathbf{d}_m^{in}$ is m-th raw materials vector, $\mathbf{x}2_m^{out}$ is m-th an evaluation vector of finished goods on model, $\mathbf{d}_m^{out}$ is m-th vector of finished goods.

7. Method of Parametrical Identification of Neural Network Model of Detection of Anomalies on the basis of Algorithm CD-1 (One-Step Contrastive Divergence)

The method of parametrical identification of neural network model of detection of anomalies on the basis of algorithm CD-1 consists of the following blocks (Fig. 3).

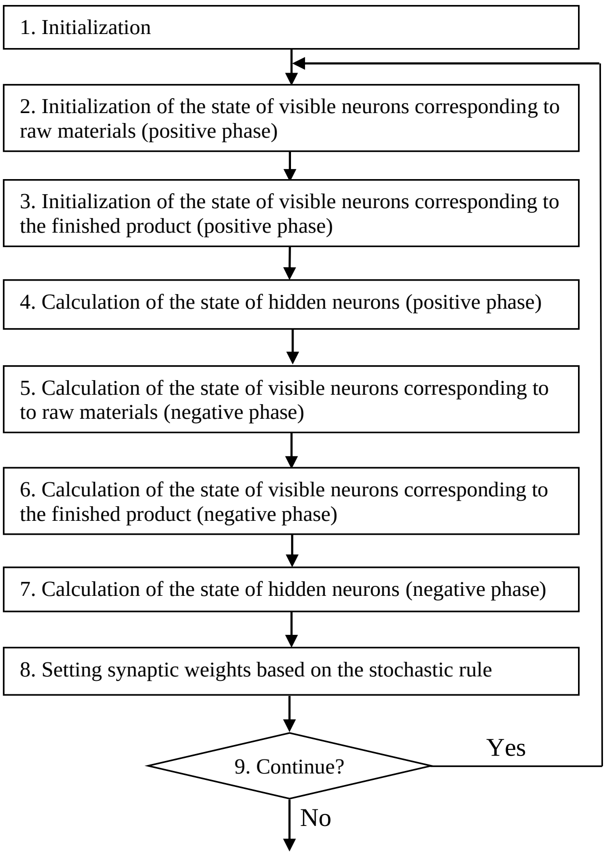

Figure 3: The sequence of procedures of a method of parametrical identification of neural network model of detection of anomalies on the basis of CD-1

1. Initialization

   Number of iteration of training $n = 1$, initialization by means of uniform distribution on an interval (0.1) or [-0.5, 0.5] bias $b_i^{in}(n)$, $i \in \overline{1, N^{in}}$, $b_i^{out}(n)$, $i \in \overline{1, N^{out}}$, $b_j^{h}(n)$, $j \in \overline{1, N^{h}}$, and weights $w_{ij}^{in-h}(n)$, $i \in \overline{1,N^{in}}$, $j \in \overline{1,N^{h}}$, $w_{ij}^{out-h}(n)$, $i \in \overline{1,N^{out}}$, $j \in \overline{1,N^{h}}$, $w_{ii}^{in-h}(n)=0$, $w_{ii}^{out-h}(n)=0$, $w_{ij}^{in-h}(n)=w_{ji}^{in-h}(n)$, $w_{ij}^{out-h}(n)=w_{ji}^{out-h}(n)$.

The training set is set $\{(\mathbf{x}_m^{in},\mathbf{x}_m^{out}) \mid \mathbf{x}_m^{in} \in (0,1)^{N^{in}}, \mathbf{x}_m^{out} \in (0,1)^{N^{out}}\}$, $m \in \overline{1,M}$, where $\mathbf{x}_m^{in}$ is m^{th} raw materials vector, $\mathbf{x}_m^{out}$ – m-th vector of finished goods, M – power of a training set, vector of mean square deviations for a raw materials vector $\boldsymbol{\sigma}^{in}=(\sigma_j^{in},...,\sigma_{N^{in}}^{in})$; vector of mean square deviations for a vector of finished goods $\boldsymbol{\sigma}^{out}=(\sigma_j^{out},...,\sigma_{N^{out}}^{out})$.

Positive phase (Step 2–4)

2. Initialization of a state of the visible neurons corresponding to raw materials $\mathbf{x}1_m^{in}=\mathbf{x}_m^{in}$, $m \in \overline{1,M}$.
3. Initialization of a state of the visible neurons corresponding to finished goods $\mathbf{x}1_m^{out}=\mathbf{x}_m^{out}$, $m \in \overline{1,M}$.
4. Calculation of a state of the hidden neurons $\left(j \in \overline{1,N^h}\right)$

$$P_{mj}=\frac{1}{2}+\frac{1}{\pi}\arctan\left(b_j^h(n)+\sum_{i=1}^{N^{in}} w_{ij}^{in-h}(n)\frac{x1_{mi}^{in}}{\sigma_i^{in}}+\sum_{i=1}^{N^{out}} w_{ij}^{out-h}(n)\frac{x1_{mi}^{out}}{\sigma_i^{out}}\right),\ m \in \overline{1,M},$$

$$x1_{mj}^{h}=\begin{cases}1, & P_{mj} \ge U(0,1)\\ 0, & P_{mj} < U(0,1)\end{cases}.\ m \in \overline{1,M}$$

Negative phase (Steps 5–7)

5. Calculation of a state of the visible neurons corresponding to raw materials $\left(j \in \overline{1,N^{in}}\right)$

$$\mu_{mj}^{in}=b_j^{in}(n)+\sigma_j^{in}\sum_{i=1}^{N^h} w_{ij}^{in-h}(n)x1_{mi}^{h},\ m \in \overline{1,M},$$

$$x2_{mj}^{in}=\mu_{mj}^{in}+\sigma_j^{in}N(0,1),\ m \in \overline{1,M}.$$

6. Calculation of a state of the visible neurons corresponding to finished goods $\left(j \in \overline{1,N^{out}}\right)$

$$\mu_{mj}^{out}=b_j^{out}(n)+\sigma_j^{out}\sum_{i=1}^{N^h} w_{ij}^{out-h}(n)x1_{mi}^{h},\ m \in \overline{1,M},$$

$$x2_{mj}^{in}=\mu_{mj}^{in}+\sigma_j^{in}N(0,1),\ m \in \overline{1,M}$$

7. Calculation of a state of the hidden neurons $\left(j \in \overline{1,N^h}\right)$

$$P_{mj}=\frac{1}{2}+\frac{1}{\pi}\arctan\left(b_j^h(n)+\sum_{i=1}^{N^{in}} w_{ij}^{in-h}(n)\frac{x2_{mi}^{in}}{\sigma_i^{in}}+\sum_{i=1}^{N^{out}} w_{ij}^{out-h}(n)\frac{x2_{mi}^{out}}{\sigma_i^{out}}\right),\ m \in \overline{1,M},$$

$$x2_{mj}^{h}=\begin{cases}1, & P_{mj} \ge U(0,1)\\ 0, & P_{mj} < U(0,1)\end{cases},\ m \in \overline{1,M}.$$

8. Setup of bias and synoptic weights on the basis of the stochastic rule

$$b_i^{in}(n+1)=b_i^{in}(n)+\eta\left(\frac{1}{M}\sum_{m=1}^{M}\frac{x1_{mi}^{in}}{\left(\sigma_i^{in}\right)^2}-\frac{1}{M}\sum_{m=1}^{M}\frac{x2_{mi}^{in}}{\left(\sigma_i^{in}\right)^2}\right),\ i \in \overline{1,N^{in}},$$

$$b_i^{out}(n+1)=b_i^{out}(n)+\eta\left(\frac{1}{M}\sum_{m=1}^{M}\frac{x1_{mi}^{out}}{\left(\sigma_i^{out}\right)^2}-\frac{1}{M}\sum_{m=1}^{M}\frac{x2_{mi}^{out}}{\left(\sigma_i^{out}\right)^2}\right),\ i \in \overline{1,N^{out}},$$

$$b_i^h(n+1) = b_i^h(n) + \eta\left(\frac{1}{M}\sum_{m=1}^{M} x1_{mi}^h - \frac{1}{M}\sum_{m=1}^{M} x2_{mi}^h\right),\ i \in \overline{1, N^h},$$

$$\rho_{ij}^+ = \frac{1}{M}\sum_{m=1}^{M} \frac{x1_{mi}^{in} x1_{mj}^h}{\sigma_i^{in}},\ i \in \overline{1, N^{in}},\ j \in \overline{1, N^h},$$

$$\rho_{ij}^- = \frac{1}{M}\sum_{m=1}^{M} \frac{x2_{mi}^{in} x2_{mj}^h}{\sigma_i^{in}},\ i \in \overline{1, N^{in}},\ j \in \overline{1, N^h},$$

$$w_{ij}^{in-h}(n+1) = w_{ij}^{in-h}(n) + \eta(\rho_{ij}^+ - \rho_{ij}^-),\ i \in \overline{1, N^{in}},\ j \in \overline{1, N^h},$$

$$\rho_{ij}^+ = \frac{1}{M}\sum_{m=1}^{M} \frac{x1_{mi}^{out} x1_{mj}^h}{\sigma_i^{out}},\ i \in \overline{1, N^{out}},\ j \in \overline{1, N^h},$$

$$\rho_{ij}^- = \frac{1}{M}\sum_{m=1}^{M} \frac{x2_{mi}^{out} x2_{mj}^h}{\sigma_i^{out}},\ i \in \overline{1, N^{out}},\ j \in \overline{1, N^h},$$

$$w_{ij}^{out-h}(n+1) = w_{ij}^{out-h}(n) + \eta(\rho_{ij}^+ - \rho_{ij}^-),\ i \in \overline{1, N^{out}},\ j \in \overline{1, N^h}$$

9. Check of termination condition

If $\frac{1}{M(N^{in}+N^{out})}\sum_{m=1}^{M}\left(\sum_{i=1}^{N^{in}} | x1_{mi}^{in} - x2_{mi}^{in} | + \sum_{i=1}^{N^{out}} | x1_{mi}^{out} - x2_{mi}^{out} |\right) > \varepsilon$ then $n = n+1$, go to 2.

8. Experiments and Results

The offered method was investigated on indicators of delivery and payment of stocks of manufacturing enterprise with a two-year depth of sample with daily time intervals.

Results of comparison of the offered neural network model (BRCM) with neural network model are presented by a bidirectional counter propagation neural network (BCPNN) on the basis of criterion of a root mean square error in Table 2.

Table 1
Comparison of the offered neural network BRCM model with traditional BCPNN on the basis of criterion of a root mean square error

Root mean square error of neural network model	
BRCM	BCPNN
0.03	0.06

According to Table 2, use of BRCM reduces a root mean square error and by that increases the accuracy of detection of anomalies

Results of comparison of the offered method of parametrical identification with use and without use of GPU and technology of parallel processing information of CUDA are provided in Table 3.

Table 2
Comparison of computing complexity of a method of parametrical identification with use and without use of GPU

Indicator	Method	
	use of GPU	without use of GPU
Computing complexity	$O(2\log_2(N^{in}N^{out}N^hM))$	$O(4(N^{in}+N^{out})N^hM)$

According to tab. 2, use of GPU reduces computing complexity approximately in $\left(2N^hM\right)/\log_2(N^hM)$ time and by that increases the speed of parametrical identification.

9. Conclusions

1. The relevant problem of increase in efficiency of detection of anomalies in data of audit of waste-free production was solved by means of neural network model of Gauss-Bernoulli bidirectional restricted Cauchy machine.
2. The proposed neural network model of Gauss-Bernoulli bidirectional restricted Cauchy machine possesses at the same time autoassociative and heteroassociative memory; real data; has no restrictions for memory capacity; provide high accuracy of anomalies detection; uses Cauchy's distribution that increases the speed of convergence of a method of parametrical identification.
3. For increase speed of parametrical identification of Gauss-Bernoulli bidirectional restricted Cauchy machine, the method of parametrical identification intended for implementation on GPU by means of CUDA technology was developed. The offered method allows to increase training speed approximately in $\left(2N^hM\right)/\log_2(N^hM)$ time, where N^h is number of neurons in the hidden layer, M is power of a training set.
4. The made experiments confirmed operability of the developed software and allow to recommend it for use in practice in a subsystem of the automated analysis of DSS of audit for anomalies detection. Prospects of further researches are in checking the offered methods on broader set of test databases.

10. References

[1] The World Bank, World Development Report 2016: Digital Dividends. URL: https://www.worldbank.org/en/publication/wdr2016.

[2] M. Schultz, M. Tropmann-Frick, Autoencoder neural networks versus external auditors: detecting unusual journal entries in financial statement audits, in: Proceedings of the 53rd Hawaii intern. conf. on System Sciences, HICSS, Maui, Hawaii, USA, 2021, pp. 5421–5430. doi: 10.24251/HICSS.2020.666.

[3] J. Nonnenmacher, et al., Using autoencoders for data-driven analysis in internal auditing, in: Proceedings of the 54th Hawaii Intern. Conf. on System Sciences, Maui, Hawaii, USA, 2021, pp. 5748–5757.

[4] T. Neskorodieva, E. Fedorov, I. Izonin, Forecast method for audit data analysis by modified liquid state machine, in: Intelligent Information Technologies & Systems of Information Security. Proceedings of the 1st Intern. Workshop, Khmelnytskyi, Jun. 10–12, 2020.

[5] T. Neskorodieva, E. Fedorov, Automatic analysis method of audit data based on neural networks mapping, Information technology and interaction, in: Proceedings of the 7th Intern. Conf., IT&I, Kyiv, Dec. 2–3, 2020.

[6] T. Neskorodieva, E. Fedorov, Method for automatic analysis of compliance of expenses data and the enterprise income by neural network model of forecast, Modern machine learning technologies and data science. Proceedings of the 2nd Intern. Workshop on MoMLeT&DS, Lviv-Shatsk, Jun. 2–3, 2020.

[7] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: A survey, ACM Computing Surveys 41(3) (2009). doi: 10.1145/1541880.1541882.

[8] V. Hodge, J. Austin, A survey of outlier detection methodologies, Artificial Intelligence Reform 22(2) (2004) 85–126.

[9] M. Agyemang, K. Barker, R. Alhajj, A comprehensive survey of numeric and symbolic outlier mining techniques, Intelligent Data Analysis 10(6) (2006) 521–538.

[10] U. P. Singh, S. Jain, A. Tiwari, R.K. Singh, Gradient evolution-based counter propagation network for approximation of noncanonical system, Soft Computing 23 (2019) 4955–4967. doi: 10.1007/s00500-018-3160-7.

[11] J.-W. Cho, H.-M. Park, Independent vector analysis followed by hmm-based feature enhancement for robust speech recognition, Sig. Process 120 (2016) 200–208.

[12] R. Chai, et al., Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system, IEEE J. Biomed. Health Inf. 21 (3) (2017) 715–724.
[13] M. I. Achmad, H. Adinugroho, A. Susanto, Cerebellar model articulation controller (CMAC) for sequential images coding, in: 2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering. doi:10.1109/ICITACEE.2014.7065734.
[14] S. Haykin, Neural Networks and Learning Machines, Upper Saddle River, Pearson Education, Inc., New Jersey, 2009.
[15] P. M. Baggenstoss, Applications of projected belief networks (PBN), in: 2019 27th European Signal Processing Conference (EUSIPCO). doi:10.23919/EUSIPCO.2019.8902708.
[16] P. Sountsov, P. Miller, Spiking neuron network Helmholtz machine. Comput. Neurosci., 2015. doi:10.3389/fncom.2015.00046.
[17] T. Kohonen, Self-organizing and associative memory, 3rd ed., Springer, New York, 2012.
[18] H. de Vries, R. Memisevic, A. Courville, Deep Learning Vector Quantization, in: ESANN 2016 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges, Belgium, Apr. 27–29, 2016.
[19] R. A. Lobo, M. E. Valle, Ensemble of binary classifiers combined using recurrent correlation associative memories, Brazilian Conference on Intelligent Systems, BRACIS, 2020, pp 442–455.
[20] M. Kobayashi, Quaternionic Hopfield neural networks with twin-multistate activation function, Neurocomputing 267 (2017) 304–310. doi:10.1016/j.neucom.2017.06.013.
[21] K. L. Du, M. N. S. Swamy, Neural Networks and Statistical Learning, Springer-Verlag, London, 2014. doi: 10.1007/978-1-4471-5571-3.
[22] Y. Park, Optimal and robust design of brain-state-in-a-box neural associative memories, Neural Networks 23(2) (2010) 210–218. doi:10.1016/j.neunet.2009.10.008.
[23] O. I. Khristodulo, A. A. Makhmutova, T. V. Sazonova, Use algorithm based at Hamming neural network method for natural objects classification, Procedia Computer Science 103 (2017) 388–395. doi:10.1016/j.procs.2017.01.126.
[24] T. Barszcz, A. Bielecki, M. Wójcik, M. Bielecka, ART-2 Artificial Neural Networks Applications for Classification of Vibration Signals and Operational States of Wind Turbines for Intelligent Monitoring, in: Dalpiaz G. et al. (Eds.), Advances in Condition Monitoring of Machinery in Non-Stationary Operations. Lecture Notes in Mechanical Engineering, Springer, Berlin, Heidelberg, 2014. doi:10.1007/978-3-642-39348-8_58.
[25] A. P. Sonika, M. Chauhan, A. Dixit, New technique for detecting fraudulent transactions using hybrid network consisting of full-counter propagation network and probabilistic network, in: 2016 International Conference on Computing, Communication and Automation (ICCCA), 2016, pp. 177–182. doi:10.1109/CCAA.2016.7813713.
[26] E. Javidmanesh, Global stability and bifurcation in delayed bidirectional associative memory neural networks with an arbitrary number of neurons, J. Dyn. Sys., Meas., Control. 139(8) (2017). doi:10.1115/1.4036229.
[27] Q. Wang, et al., A novel restricted boltzmann machine training algorithm with fast Gibbs sampling policy, Mathematical Problems in Engineering, 2020. doi:10.1155/2020/4206457.
[28] V. Buriachok, et al., Invasion Detection Model using Two-Stage Criterion of Detection of Network Anomalies, Cybersecurity Providing in Information and Telecommunication Systems (CPITS), pp. 23–32, Jul. 2020.
[29] A. Carlsson, et al. Sustainability Research of the Secure Wireless Communication System with Channel Reservation. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 2020. https://doi.org/10.1109/tcset49122.2020.235583
[30] V. Lakhno, et al., Funding Model for Port Information System Cyber Security Facilities with Incomplete Hacker Information Available, Journal of Theoretical and Applied Information Technology 96(13), 4215–4225, 2018.