UDC 004.9
DOI: 10.15587/2312-8372.2018.135483

Bidyuk P.,
Prosyankina-Zharova T.,
Terentiev O.,
Medvedieva M.

ADAPTIVE MODELLING FOR FORECASTING ECONOMIC AND FINANCIAL RISKS UNDER UNCERTAINTY CONDITIONS IN TERMS OF THE ECONOMIC CRISIS AND SOCIAL THREATS

Об'єктом дослідження є соціально-економічні процеси в контексті структурних перетворень, що відбуваються внаслідок суспільно-політичної кризи в країні. Одним з найбільш проблемних місць є відсутність комплексного дослідження та обґрунтування застосування інструментів прогнозування потенційних загроз у гуманітарній і соціальній сферах та визначення шляхів їх подолання, спрямованих на стабільний та позитивний розвиток національної економіки.

В ході дослідження використовувались системний аналіз та елементи теорії систем, методи математичного та економетричного моделювання. Системний аналіз та теорія систем використовуються для вивчення стану та поведінки національної економіки та її підсистем в сучасних умовах невизначеностей та ризиків, характерних для соціальних потрясінь та структурних змін. Методи математичного і статистичного моделювання та теорії прийняття рішень були використані для прогнозування розвитку нестаціонарних нелінійних процесів, що характерні для сучасної української економіки.

Розглянуто проблему розробки методів вирішення задач моделювання та оцінки окремих типів ризиків з можливістю застосування альтернативних методів обробки даних, моделювання та оцінки параметрів і станів національної економіки та її складових у сучасних умовах суспільно-політичних перетворень та структурних реформ. Для того, щоб знайти «найкращу» структуру моделі, рекомендується застосувати адаптивні схеми оцінювання, які передбачають автоматичний пошук у визначеному діапазоні параметрів структури моделі (тип розподілу, зменшення розмірності моделі, часові лаги та нелінійності). Запропоновані схеми адаптивної оцінки також допомагають розкрити структурні та параметричні невизначеності. Запропонована загальна методологія призначена для вирішення обраної проблеми прогнозування динамічних процесів та оцінювання кількох видів соціально-економічних та фінансових ризиків з використанням відповідних статистичних даних в комп'ютерних системах підтримки прийняття рішень.

Результати дослідження будуть корисними і для інших країн, де відбуваються аналогічні процеси.

Ключові слова: *адаптивне моделювання, ідентифікація невизначеності, оцінка ризиків, система підтримки прийняття рішень.*

1. Introduction

Today in a complex socio-political and economic situation in terms of growing influence of external factors, presence of uncertainties and risks there exists a problem of anticipating potential threats in the humanitarian and social spheres.

The problems of preventing losses and providing response to potential risks, promoting the growth of national economy and the welfare of the citizens and not to worsen the ecology situation are complex and characterized of the different uncertainties. The complexity of decisions regarding these tasks is in the presence of significant amounts of quantitative and qualitative information, various uncertainties, and existence of complex causal relationships between the factors.

To solve these problems, the methods of choice and justification of specific techniques to solving selected problem of mathematical modelling and forecasting dynamic processes under study are selected and applied. Forecasting and estimation of several kinds of risks for socio-economic systems using appropriate statistical data, and provide examples demonstrating how these methods could be applied in decision support systems were performed. The proposed methods provide an integrated analysis of factors and risks on the development of the situation in different spheres of national economy. That is why all the tasks of this study are very up-to-date.

2. The object of research and its technological audit

The object of research is nonlinear and nonstationary economic and financial processes taking place in conditions of uncertainties and risks that take place under the influence of structural transformations as a result of socio-political crisis in the country.

Development of national economy in the context of reforms, as noted by many authors [1–6], is accompanied by various uncertainties which increase the probability of risks arising from the adoption of wrong control decision. Therefore proceeding from the position of the general theory of systems [1, 7, 8], and given that the socio-economic system has a hierarchical structure, its information model can be represented as follows:

$$S_0 = S_1 \cdot S_2 \cdot ... \cdot S_i \cdot ... S_m, \quad (1)$$

where S_i is i-th hierarchical level; m is a number of hierarchical levels; and:

$$S_i = \langle M_i, P_i, R_i, X_i, Y_i, f_i, \varphi_i \rangle, \quad (2)$$

where M_i, P_i, R_i are the sets of real objects, subjects and subsystems on i-th hierarchical level; X_i, Y_i are the sets of internal and external parameters of the system of the i-th hierarchical level and external environment; φ_i, f_i are the functions, that determine the relationship of appropriate parameters on m levels as follows:

$$\varphi_i : X_i \to Y_i;\ f_i : Y_i \to Y_{i-1}. \quad (3)$$

Starting from the fact that the problem under study is complex and weakly structured, and the studied system is complex hierarchical, this research is performed according to the scheme presented in Table 1.

Table 1

The sequence of stages for predictive modelling of risks for socio-economic and financial processes using an adaptive approach

Phase and number of the stage and actions performed on this stage	Stage requirements
1. Collection and processing of statistical data of socio-economic indicators	Organization of collecting the information from different sources, providing uniformity of files and forms for information collecting; clarity, precision, uniqueness and completeness of data, preliminary verification of quality of gathered data at the place of their collection
2. Formation of databases	Providing security and confidence for data in the databases, analysis of efficiency and reliability of data transmission
3. Pre-processing, formalization and structuring of data sets	Data quality testing (completeness, homogeneity, information content), formalization of data in the system (standardization of presentation forms). Alignment with requirements of data analysis methods, forming of information repository for centralized processing
4. Statistical analysis of data, model construction, estimation of forecasts and possible risks of losses	Ensuring uniqueness of interpretation of the results obtained, analyzing quality of the constructed models and the quality of estimates of forecasts using a set of relevant statistical characteristics of quality. Selecting the best method for evaluating the structure and parameters of mathematical models (ordinary least squares, maximum likelihood estimation, Markov chain, Monte Carlo)
5. Analysis of results; risk management, providing recommendations for possible improvement of current results	On the basis of the constructed models for estimating the forecast for subsequent periods, establishing of objective causes and factors influencing the results. Studying by analytics and managers the results, checking their quality and reliability, using the results obtained for the assessment and risk management, and generating recommendations for application them in further research

As it is represented in the Table 1, the uncertainties are the factors that influence negatively over all steps of the data processing and predictive modelling, risk estimation and decisions generating directed towards minimization of possible loss. In many cases, the researchers [8, 9] have to cope with general types of uncertainties.

In addition, the need for handling uncertainties and risks of a different nature requires more active use of system tools in the decision-making process. Experience shows [8–11] that the best results for assessing the risks of processes are usually achieved using ideologically different methods, combined within a single computer system.

3. The aim and objectives of research

The aim of research is increasing the quality of predictive modelling of losses due to economic and financial risks in conditions of economic and social threats.

To achieve the aim the following tasks were solved:

1. Choice and justification of applied techniques to solving selected problem of mathematical modelling and forecasting dynamic processes under study.
2. Forecasting and estimating several kinds of risks for socio-economic systems using appropriate statistical data.
3. Development of methodological bases for adaptive algorithm that provides a possibility for easy extension by the new parameter estimation techniques, forecasting methods, economy and financial risks estimation procedures as well as decision alternatives generation.
4. Providing examples demonstrating how these methods could be applied in decision support systems.

4. Research of existing solutions of the problem

Many researches [2–6] are devoted to development of methodological bases of predictive modelling for economic and financial processes to the use of different approaches. In the overwhelming majority of studies [7–11], the problems are considered in conditions of stable economic and socio-political situation.

However, as noted in [1–5], there are significant differences in modelling of these processes under influence of various types of risks and uncertainties in terms of structural changes and by the stable economy. It is noted [1–4, 11] in condition of non-stable economic and socio-political situation, economic and financial processes are not only non-stationary and non-linear, characterized by seasonality, but also there is a high probability of occurrence of various types of risks. Also, the authors [2–6], noted, in this terms, the previous state of the system cannot clearly describe its condition and behaviour in the future. That is why, the authors of [1] state that under such conditions, the use of one, even well-developed methodology in conditions of a stable economy, is not sufficient for obtaining high quality forecast. The authors of [9–12] have shown that under such conditions it is practically impossible to form long enough time series of statistical and other investigated indicators, there is a risk of presence of a large number of data breaks, extreme values, non-comparable indicators, external noise etc. In addition, the data obtained through the survey are often noisy, and contain gaps. And it is necessary to use preliminary processing of input data.

 ISSN 2226-3780

So there is a need for developing such method of predictive modelling, which would allow creation of system models with given parameters. To solve all described tasks it should be used a systemic interdisciplinary approach, system analysis methodology and adaptive modelling. Good opportunities for effective satisfaction of all requirements are given by the use of these methods in a single computing system, as noted by [9, 13] that provides for enhancing the overall behavioural effect for specific applications in the field of constructing models with statistical data.

Summing up it should be noted that the major differences over the use of different approaches to predictive modelling described in scientific literature can be grouped as follows:

– application of systemic approach to the study of socio-economic systems;
– construction of system models with given parameters;
– application of adaptive modelling method;
– use of DSS with specialized functions that implement adaptive modelling techniques and lost data imputation.

Thus, the results of analysis of information sources suggest that implementation of adaptive modelling methodology for economic and financial processes will not only improve quality of the forecasts. Its use in DSS [13] will allow creation of universal, highly developed decision support system with hierarchical architecture compatible with the nature of human decision-making.

5. Methods of the research

Some types of economic risks are estimated with different modifications of well-known Value-at-Risk (VaR) methodology that provides a possibility for reaching practically acceptable quality of risk estimates [8–12]. To analyse financial risks the scoring models are widely used as of today: linear and nonlinear regression, Bayesian networks (BN), decision trees, fuzzy logic, factor analysis, support vector machine (SVM), neural and neuro-fuzzy techniques, as well as combinations of the approaches mentioned [7, 8, 11, 12]. But in conditions of economic reforms and social crisis better results could be achieved by the usage of adaptive estimation scheme [13]. To find «the best» model structure, as it is recommended by [8, 13], it is necessary to apply adaptive estimation schemes that provide for an automatic search of model structure and its parameters.

The adaptive estimation schemes also help to cope with the model parameters uncertainties. New data are used to compute repeatedly model parameter estimates that correspond to possible changes in the object under study. In the cases when model is nonlinear alternative parameter estimation techniques could be hired to compute alternative (though admissible) sets of parameters and to select the most suitable model using a set of statistical quality criteria (Fig. 1).

To achieve reliable high quality final result of risk estimation and forecasting at each stage of computational hierarchy separate sets of statistical quality criteria have been used. Data quality control is performed with the following criteria: number of missing values; testing for outliers; data normalization and filtering; computing extra indicators.

The procedure for assessing risks and uncertainties has a situational iterative nature. Choice of different options for the situation development and the degree of risk can be interpreted by the decision making person in different ways. The socio-economic process at the stage of constructing mathematical models for predictive modelling is advisable to present in the form of a model in the space of states [4]:

$$x(k) = A(k,k-1)x(k-1) + B(k,k-d)u(k-1) + w(k),$$
$$z(k) = Hx(k) + v(k), \qquad (4)$$

where $x(k)$ is n-dimensional vector of system states; $k = 0, 1, 2, \ldots$ is discrete time; $u(k-1)$ is m-dimensional vector of deterministic control variables; $w(k)$ is n-dimensional vector of external random disturbances; $A(k, k-1)$ is $(n \times n)$ matrix of system dynamics; $B(k, k-1)$ is $(m \times n)$ matrix of control coefficients.

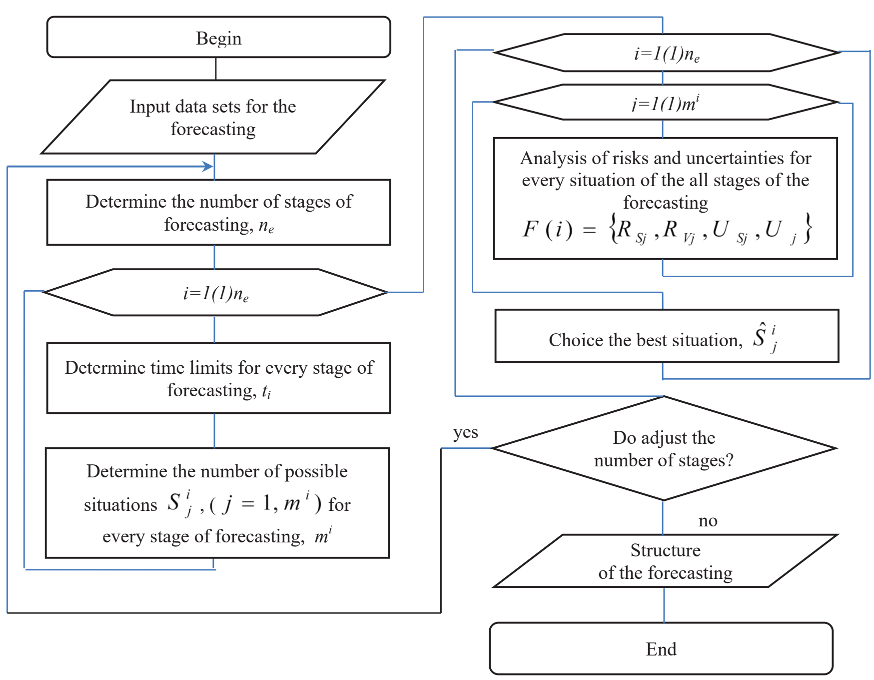

Fig. 1. The procedure for adaptive model structure estimation

The double argument $(k, k-1)$ – means that the variable or parameter is used at the moment k, but its value is based on the former (earlier) data processing including:

– moment $(k-1)$;
– discrete time k and continuous time t are linked to each other via data sampling time T_s: $t = k\,T_s$;
– d is delay time of the system under study;
– $z(k)$ is vector of measurement of r elements;
– $H(k)$ is matrix of coefficient of measurement (very often it has a diagonal form);
– $v(k)$ is vector of noise components (measurement errors).

In order to reduce uncertainty in the form of influence of two random processes $w(k)$ and $v(k)$, and to evaluate non-measurable components of state vector, it is expedient to use optimal Kalman filter [14]. In addition, in forecasting the risks of socio-economic processes, for repeated evaluation of the system matrix, and covariance of two random processes, it is better to use adaptive versions of the filters.

The unified notion of a model structure is as follows [13]:

$$S = \{r, p, m, n, w, l\}, \qquad (5)$$

where r is model dimensionality (number of equations that constitute the model); p is model order (maximum order of differential or difference equation in a model); m is a number of independent variables in the right hand side of a model; n is a nonlinearity and its type; w is stochastic external disturbance and its type; l are possible restrictions for the variables and/or parameters.

In this case, model structure should practically always be estimated using data. When a model is constructed for forecasting let's build several candidates and select the best one of them using a set of model quality statistics. To select the best model constructed the following statistical criteria are used: determination coefficient (R^2), Durbin-Watson statistic (DW), Fisher F-statistic, Akaike information criterion (AIC) and residual sum of squares (RSS).

The forecasts quality is estimated with hiring the criteria mean squared error (MSE), mean absolute percentage error ($MAPE$), Theil inequality coefficient (U).

To perform automatic model selection the combined criterion could be hired. The power of the criterion was tested experimentally and proved with a wide set of models and statistical data. Thus, the three sets of quality criteria are used to insure high quality of final result. So, the adaptive estimation scheme for building the «best» model is based on the criterion:

$$V_N(\theta, D_N) = e^{|1-R^2|} + e^{|2-DW|} + \alpha \ln\left(1 + \frac{SSE}{N}\right) +$$
$$+ \beta\left\{\ln(1 + MSE) + \ln(1 + MAPE)\right\},$$

where θ is a vector of model parameters; D_N is data in the form of time series (N is a power of time series used); R^2 is a determination coefficient; DW is Durbin-Watson statistic; MSE is mean square error; $MAPE$ is mean absolute percentage error for forecasts; α, β are adjustment coefficients that could be selected by a user or searched automatically by the decision support system itself.

6. Research results

The methodology of data imputation in adaptive scheme as proposed above was applied for short term forecasting of agriculture development in Ukraine. The indices of production of the main agricultural products in the period within 1940–2014 [15] were used as the experimental data. The forecasting quality was tested with the data related to 2012–2014. About 10 % of input data is missing. The method was proposed for data imputation based on the use of the five available previous measurements weighted by the coefficients selected in a special way. Several of the coefficients were computed by the method analogous to exponential smoothing.

Development of this sphere is represented by production volume of the main agricultural crops. Using the adaptive approach proposed statistical analysis of data, model construction, estimation of forecasts and possible risks of losses were provided. In computational experiments observed the 11 agricultural time series and were performed the steps given below:

1. Generating 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 % of missing data in the dataset.

2. For each variable were created ten samples each of them contained missing values with the appropriate portion.

3. Using various methods of imputation were the missing values filled. The best of them are follows:

Method 1. Usage of standard linear regression $X_1 = a_0 + a_2 \cdot X_2 + \ldots + a_{11} \cdot X_{11}$.

Method 2. Replacing by the average value computed with the sample.

Method 3. Imputation of missing values by the values calculated for each missing value as an average between the first previous and the first subsequent values.

Method 4. For each missing value, based on the five previous values, the mean is calculated using weighting coefficients according to the principle of exponential smoothing.

For example, if X_1 contains missing value at time t, then the substitution value will be calculated using the formula:

$$X_1(t) = 0.5 \cdot X_1(t-1) + 0.25 \cdot X_1(t-2) +$$
$$+ 0.125 \cdot X_1(t-3) + 0.1 \cdot X_1(t-4) + 0.025 \cdot X_1(t-5).$$

The weighting coefficients were set using the method of exponential smoothing.

Method 5. Using the model of autoregression (*AR(p)*) based on the assumption that there exists 3-year cycles in the data-seasonality.

General results for missing data imputation are presented in Table 2.

As comparative analysis of using different methods of imputation showed, the best results were achieved using linear regression. This method could be used for any percentage of missing data. The methods of missing values imputation by the average between the first previous and the first subsequent values, usage the exponential smoothing of 5 previous values, usage of autoregression model could not be used when the number of missing data was more than 20, 25 and 15 % respectively, because in those cases it would appear long sequences of omissions.

 ISSN 2226-3780

Table 2

The means of *MAPE* which were computed by predictive modelling of agriculture indicators using various methods of imputation for missing data

Main indicators of agricultural production	Methods of missing values imputation				
	Using the standard linear regression	Replacing using the average value by the sample	Impute missing value by the average between the first previous and the first subsequent values	Exponential smoothing	Using the model of auto-regression
Grain and leguminous crops (x_1)	15.055	32.182	13.41	19.639	20.943
Sugar beet (factory) (x_2)	19.094	69.172	15.264	23.900	30.368
Sunflower (x_3)	32.139	78.297	11.274	19.239	22.738
Potatoes (x_4)	11.141	13.534	8.373	12.738	12.731
Vegetables (x_5)	10.116	28.780	6.504	10.931	12.644
Meat (in slaughter weight) (x_6)	8.406	46.918	6.319	14.716	12.007
Milk (x_7)	5.537	35.378	4.761	9.166	9.079
Beef and veal (x_8)	6.846	66.220	5.552	12.778	10.085
Cattle (x_9)	4.479	42.590	3.743	8.3819	7.105
Pigs (x_{10})	11.028	54.886	10.190	16.366	21.437
Sheeps and goats (x_{11})	22.817	97.532	7.342	17.132	15.211
Average *MAPE*	12.535	45.897	8.370	14.610	15.300

Short term forecasting for the gross domestic product (*GDP*) in Ukraine in terms of structural reforms is considered. Among possible independent variables are the following: consumer price index (*CPI*), production price index (*PPI*), tax deductions (*TD*), and natural gas price (*NGP*).

The model types, shown in Table 3, are as follows: autoregression with different lags, autoregression $AR(12)$ with linear trend ($AR(12)+t1$), autoregression $AR(12)$ with explaining variables such as *CPI, PPI, TD* and *NGP*. Dynamic Bayesian network (*BN*) was constructed using the same variables. The best forecasting result provided $AR(12)$ with linear trend that resulted with $MAPE = 4.57$ %.

Table 3

Results of forecasting model building for *GDP*

Model type	Model adequacy		Forecast quality (*MAPE*)	
	R^2	DW	traditional	adaptive algorithm
AR(1)	0.901	1.87	13.99	12.78
AR(5)	0.924	2.04	10.37	9.03
AR(12)	0.978	1.63	5.69	4.85
AR(12)+*t*1	0.982	1.94	4.74	4.57
AR(12)+*CPI*+*PPI*+*TD*+*NGP*	0.981	1.57	5.51	5.38
Dynamic *BN*	–	–	–	6.49

The problem of credit borrowers' classification or their solvency estimation is considered as another example. The following characteristics of clients were used: previous solvency data, age, credit sum, type of currency (local or USD), term of crediting, term of residing in defined area, marital status, number of children, type of credit, and gender. The uncertainty met in this example was in the form of incomplete records. It was established with Bayesian network that maximum model accuracy reached was 0.764 at the cut-off value 0.3. The statistical criteria characterizing quality of the models constructed are given in Table 4.

The best models for estimation of credit return probability turned out to be discriminant analysis, logistic regression and Bayesian network in both cases (the use of commercial software and the *DSS* proposed). The best common accuracy (*CA*) was shown by discriminant analysis (0.837 and 0.865), and logistic regression (0.798 and 0.829), though Bayesian network showed higher Gini index than logistic regression (0.689). The decision tree used is characterized by Gini index of about 0.583, and $CA = 0.763$. It should be stressed that acceptable values of Gini index for developing countries like Ukraine are in the range 0.4–0.6.

Table 4

Quality of the models constructed for clients solvency estimation

Model type	Common method		Common accuracy	
	Gini index	AUC	common method	proposed method
Discriminant analysis	0.723	0.891	0.837	0.865
Bayesian network	0.689	0.845	0.764	0.787
Logistic regression	0.678	0.847	0.798	0.829
Decision tree (*CHAID*)	0.583	0.791	0.763	0.774
Linear regression	0.386	0.647	0.616	0.637

7. SWOT analysis of research results

Strengths of the method are:
– increase of the quality of the forecast estimates creates conditions for optimizing decision-making process regarding control of particular industries development and the national economy as a whole;
– achieving the best results of the usage of the ideologically different techniques of modelling and risk forecasting what creates a convenient basis for combination of various approaches;
– estimation schemes proposed also help to cope with the model parameters uncertainties;
– the method was proposed for data imputation that is effective even if there are 30 percent of gaps.

Weaknesses are:

– lack of instant results due to appropriate tracking of the computational processes at all data processing stages;

– it is necessary to prepare carefully the input data sets;

– the experts should be involved at the stage of preparation of data due to the need to process time series with a significant number of gaps.

Opportunities are:

– rapid adaptation with the use of new parameters, estimation techniques, forecasting methods;

– automatic model selection realization is very easy in modern decision support systems;

– quality increasing of results of prognostic modelling is achieved thanks to appropriate tracking of the computational processes at all data processing stages.

In the future studies it is supposed to expand DSS functionality with new model types, namely combinations of statistical and intellectual data analysis techniques, as well as new parameter estimation methods enhancing DSS possibilities regarding various data distributions. The criteria bases will also be expanded with improved combined criteria for automatic model selection. The whole data processing procedure including model building, forecast estimating and alternative decisions generating will be controlled automatically by DSS itself.

Threats are:

– risks of wrong decision due to lack of experience in decision-making in the context of sharp socio-political changes in the country;

– lack of long enough time series of socio-economic and financial indicators for processing;

– risk does not take into consideration all factors relevant to the specific study presented.

8. Conclusions

1. The study of adaptive modelling methods for forecasting development of economic and financial processes and prediction of losses of the risks under uncertainty is carried out. The general methodology for mathematical modelling and forecasting economic and financial processes, and risk estimation that is based on system analysis principles is considered. Methodology for application of optimal Kalman filter, multiple missing data imputation techniques, alternative methods for model parameter estimation, and Bayesian programming approach of applied techniques to solving selected problem was justified.

2. In current circumstances of reforms and abrupt changes in macroeconomic policy using appropriate statistical data in the context, the decision maker faces the problem simultaneously to overcome the uncertainties of almost all types, namely structural, statistical, parametric, probabilistic and amplitude uncertainty. That is why it was proposed the scheme of the research in the form of predictive modelling the risks for socio-economic and financial processes using adaptive approach.

3. Also the method was proposed based on adaptive algorithm that provides a possibility for easy extension by the new parameter estimation techniques, forecasting methods, economy and financial risks estimation procedures as well as decision alternatives generation. The adaptive estimation schemes that also help to cope with the model parameters uncertainties were proposed. The feature of this adaptive algorithm is that it provides a possibility for easy extension of its functional possibilities with new parameter estimation techniques, forecasting methods, economic and financial risks estimation procedures, and decision alternatives generation. High quality of the final result is achieved thanks to appropriate tracking of computational processes at all data processing stages: preliminary data processing, model structure and parameter estimation, computing of short- and middle-term forecasts, and estimation of risk variables (parameters) as well as thanks to convenient for a user intermediate and final results representation. The modelling methodology is based on ideologically different techniques of modelling and risk forecasting what creates a convenient basis for combination of various approaches to achieve the best results.

4. The methods proposed in the frames of decision support system were used successfully for solving practical problems of dynamic processes forecasting and risk estimation in economic and financial systems. The methodology of data imputation in adaptive scheme as proposed above was applied to short term forecasting non-stationary non-linear processes characterized by seasonality. While solving this task the method was proposed for data imputation based on the use of the five available previous measurements weighted by specially selected coefficients. As another example considered was the problem of credit borrowers' classification or their solvency estimation. The best models for estimation of the credit return probability turned out to be discriminant analysis, logistic regression and Bayesian network in both cases (the combined use of commercial software and the *DSS* proposed).

Application of the proposed methodology will allow promoting the growth of national economy, ensuring the national food security, controllability of ecological situation etc.

References

1. Zgurovsky M. Z. Technology foresight of ukrainian economy in the medium (up to 2020) and long term (until 2030) time horizons // Visnik Nacional'noi' Akademii' Nauk Ukrai'ni. 2016. No. 1. P. 57–68. doi: http://doi.org/10.15407/visn2016.01.057
2. Szeles M. R., Munoz R. M. Analyzing the regional economic convergence in Ecuador. Insights from parametric and nonparametric models // Romanian Journal of Economic Forecasting. 2016. Vol. XIX (2). P. 43–65.
3. Monetary effectiveness in small transition economy – the case of the republic of Serbia / Zivkov D. et al. // Romanian Journal of Economic Forecasting. 2016. Vol. XIX (3). P. 5–18.
4. Garcia M. G. P., Medeiros M. C., Vasconcelos G. F. R. Real-time inflation forecasting with high-dimensional models: The case of Brazil // International Journal of Forecasting. 2017. Vol. 33, No. 3. P. 679–693. doi: http://doi.org/10.1016/j.ijforecast.2017.02.002
5. Stovicek K. Forecasting with ARMA models. The case of Slovenian inflation // Prikazi in analize. 2007. Vol. XIV/1. P. 23–55.
6. Dobrescu E. Modelling an Emergent Economy and Parameter Instability Problem // Journal for Economic Forecasting of Institute for Economic Forecasting. 2017. Vol. 2. P. 5–28.
7. Arora S., Little M. A., McSharry P. E. Nonlinear and non-parametric modeling approaches for probabilistic forecasting of the US gross national product // Studies in Nonlinear Dynamics and Econometrics. 2013. Vol. 17, No. 4. P. 395–420. doi: http://doi.org/10.1515/snde-2012-0029
8. Glass K. Real-time Macroeconomic Data and Uncertainty // DEP (Socioeconomics) Discussion Papers. Macroeconomics and Finance Series. 2014. Vol. 6. P. 1–29.

 ISSN 2226-3780

9. Eeckhoudt L., Gollier C., Schlesinger H. Economic and financial decisions under uncertainty. Princeton: Princeton University Press, 2004. 248 p.
10. Penman S. H. Financial Forecasting, Risk and Valuation: Accounting for the Future. Abacus. 2010. Vol. 46, No. 2. P. 211–228. doi: http://doi.org/10.1111/j.1467-6281.2010.00316.x
11. Clements M. P., Franses P. H., Swanson N. R. Forecasting economic and financial time-series with non-linear models // International Journal of Forecasting. 2004. Vol. 20, No. 2. P. 169–183. doi: http://doi.org/10.1016/j.ijforecast.2003.10.004
12. Sekerke M. Bayesian Risk Management: A Guide To Model Risk And Sequential Learning In Financial Markets. Hoboken: John Wiley & Sons, Inc., 2015. 240 p. doi: http://doi.org/10.1002/9781118864784
13. Bidyuk P., Gozhyj O., Korshevnyuk L. Development of Decision Support Systems. Mykolaiv: Petro Mogyla Black Sea National University, 2012. 379 p.
14. Petersen I. R., Savkin A. V. Robust Kalman filtering for signals and systems with large uncertainties. Boston: Birkhouser, 1999. 202 p. doi: http://doi.org/10.1007/978-1-4612-1594-3
15. Statistical publication // The official website of the State Statistics Service of Ukraine. 2016. URL: http://www.ukrstat.gov.ua/

Bidyuk Petro, *Doctor of Technical Sciences, Professor, Department of Mathematical Methods of Systems Analysis, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Ukraine, ORCID: https://orcid.org/0000-0002-7421-3565, e-mail: pbidyuke_00@ukr.net*

Prosyankina-Zharova Tatyana, *PhD, Department of Physical and Mathematical Modelling, Institute of Telecommunications and Global Information Space, Kyiv, Ukraine, ORCID: https://orcid.org/0000-0002-9623-8771, e-mail: t.pruman@gmail.com*

Terentiev Oleksandr, *PhD, Junior Researcher, Department of Mathematical Methods of Systems Analysis, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Ukraine, ORCID: https://orcid.org/0000-0002-4288-1753, e-mail: o.terentiev@gmail.com*

Medvedieva Mariia, *PhD, Associate Professor, Head of the Department, Department of Informatics and Information and Communication Technologies, Pavlo Tychyna Uman State Pedagogical Universit, Uman, Cherkasy region, Ukraine, ORCID: https://orcid.org/0000-0001-9330-5185, e-mail: medvedeva-masha25@ukr.net*

UDC 004.915
DOI: 10.15587/2312-8372.2018.141299

Petrenko A., Kyslyi R., Pysmennyi I.

DESIGNING SECURITY OF PERSONAL DATA IN DISTRIBUTED HEALTH CARE PLATFORM

Об'єктом дослідження є розробка системи електронної медкарти (EHR), призначеної одночасно як для взаємодії пацієнт-лікуючий лікар, так і для обміну анонімізованими даними між різними медичними організаціями для їх подальшої обробки та побудови аналітичних моделей. Постійний моніторинг стану пацієнта, а також кількість і якість оброблених даних є ключовими факторами, що впливають на точність постановки діагнозу і подальші лікарські рекомендації. Слід зауважити, що більшість сучасних підходів до проектування EHR-систем є вразливими до атак цілісності даних і не дозволяють обмінюватися інформацією з іншими організаціями, зберігаючи при цьому лікарську таємницю, що призводить до наявності у окремих акторів лише невеликих фрагментованих датасетів. Важливим напрямком для поліпшення існуючих рішень є безпека обміну інформацією між натільними смарт-сенсорами.

У даній роботі пропонується розбити архітектуру на шари з виділеними зонами безпеки. Ця фрагментація дозволяє ефективно сегментувати інфраструктуру, дозволяючи кожному елементу застосовувати свої власні вимоги до аутентифікації і авторизації, використовуючи різні підходи до захисту інформації. Додатковим ефектом цього підходу є зниження навантаження на мережу і уникнення проблем безпеки шляхом мінімізації передачі конфіденційних даних (наприклад, проводити базовий збір та обробку даних на смарт-сенсорах). Пропонується використання блокчейн-технологій для забезпечення цілісності даних з використанням офф-чейн бази даних для оптимізації зберігання та швидкості транзакцій. Застосування MPC-протоколу дозволяє обмінюватися даними між партнерськими організаціями для спільних розрахунків і навчання ml-моделей, не показуючи фактичні дані.

Пропоновані підходи дозволяють створювати надійну, гнучку і в той же час безпечну платформу для збору конфіденційних даних, їх аналізу і обробки розподіленою багатоакторною системою, використовуючи переваги туманних обчислень, блокчейна та MPC.

Ключові слова: *безпека електронної карти пацієнта, блокчейн в медицині, безпека особистих даних, безпека мереж натільних датчиків.*

1. Introduction

Fast development of different wearable health monitoring and tracking capable devices allows collection and processing of patient's data during his everyday activity. By integrating these data sources with Personal Health Record (PHR) systems and partly automating inferences and decision-making process for physician assistance that can be applied to the new connected e-Health paradigm. It allows simultaneous integration of PHRs information, body sensor networks' (BSN) streams, activity data and context for better outcomes in preventive health as well as

