

**О ПРОГНОЗИРОВАНИИ ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ
С ПОМОЩЬЮ НЕЙРОЭВОЛЮЦИОННЫХ МОДЕЛЕЙ**

Д. В. Федотов, Е. С. Семенкин

Сибирский государственный аэрокосмический университет имени академика М. Ф. Решетнева
Российская Федерация, 660014, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: fedotov.dm.v@gmail.com, eugenesemenkin@yandex.ru

Прогнозирование временных рядов является сложной и интересной задачей. Экономическая ситуация в стране или отрасли зависит от многих факторов и событий, а также, в свою очередь, влияет на множество ситуаций. Инновационные проекты в области ракетно-космической техники требуют больших денежных инвестиций и временных затрат на разработку и внедрение. Для таких вложений необходима релевантная и надежная информация о возможном будущем состоянии экономики с целью минимизации рисков, а также для своевременной корректировки планов. В частности, для получения такой информации можно использовать временные ряды различных показателей экономического развития. В данной статье в качестве основной технологии прогнозирования используются нейросетевые предикторы, автоматически генерируемые генетическими алгоритмами, что позволяет повысить качество предоставляемого прогноза, а также показывается преимущество самоконфигурируемого генетического алгоритма по сравнению со стандартным при использовании в программной системе для настройки весовых коэффициентов нейронных сетей. Качество прогноза, выполняемого нейронной сетью, зависит также и от ее структуры. В работе используется алгоритм генетического программирования для формирования структуры нейронных сетей, так как он позволяет создавать гибкие и эффективные модели без участия человека-эксперта, что позволяет автоматизировать конфигурирование нейросетевых предикторов. Тестирование полученной комбинации интеллектуальных информационных технологий осуществляется на задаче прогнозирования значений финансовых временных рядов с использованием в качестве входных значений цен и объемов торгов, а также полученных на их основе технических индикаторов. Разработанная система позволяет автоматически построить нейросетевой предиктор и получить качественный прогноз изменения экономических показателей. Проведено сравнение с другими интеллектуальными информационными технологиями прогнозирования и выявлено, что разработанная система показывает в среднем более высокие результаты.

Ключевые слова: нейронные сети, эволюционные алгоритмы, самоконфигурирование, временные ряды, прогнозирование.

Vestnik SibGAU
2014, No. 5(57), P. 299–304**ON FORECASTING ECONOMIC INDEXES
BY MEANS OF NEUROEVOLUTIONARY MODELS**

D. V. Fedotov, E. S. Semenkin

Siberian State Aerospace University named after academician M. F. Reshetnev
31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660014, Russian Federation
E-mail: fedotov.dm.v@gmail.com, eugenesemenkin@yandex.ru

Time series prediction is quite complex and interesting problem. Usually a forecast is based on information from the past, i.e. on previous values of features. The economic situation in the country depends on different factors and events and influences many situations. The innovation projects in aerospace field require a large amount of investments and take a long time for development and implementation. These kinds of investments require relevant and reliable information about possible future economic situation to minimize risks and correct the plans in right time. In particular, to get such information one may use the time series of the various indexes of economic development. In this paper, the neural network predictors are used as the main technology of forecasting. They are generated automatically by means of genetic algorithms that allow increasing the quality of the prediction. Besides, the advantages of the use of self-configuring genetic algorithms in comparison with standard genetic algorithms are demonstrated when adjusting the weighting coefficients of the neural network. The quality of the solution provided by the neural network depends also on its structure. In this paper, the genetic programming is used for the neural network structure design due to its ability to provide effective and flexible models without a human expert. This allows automating the neural network based predic-

tors configuring. The combination of these intellectual information technologies is tested on the financial time series forecasting task using the values of the prices and volumes along with the technical indicators as the inputs. The developed system allows automated designing the neural network based predictors and getting a high quality forecast of economic indexes changes. The developed system shows better results in the comparison with the other forecasting technologies.

Keywords: neural network, evolutionary algorithms, self-configuring, time series, forecasting.

Введение. Современный мир характеризуется динамично развивающейся экономикой. Свое влияние на нее оказывают события в различных областях: политической, научной и др. За последние 20 лет российская экономика несколько раз была подвержена кризисам. Инновационные проекты в области ракетно-космической техники требуют привлечения больших объемов инвестиций, а также занимают длительное время на разработку и внедрение. Подобные вложения требуют осторожности и предусмотрительности, в частности, тщательного прогнозирования развития экономики и финансовой системы как страны в целом, так и ее ракетно-космической отрасли. Также необходимо эффективно и своевременно реагировать на возможные изменения показателей экономической динамики, производить коррекцию производственных планов и механизмов финансирования проектов. Для этого следует использовать интеллектуальные предикторы с возможностью их автоматического перепроектирования в оперативном (автоматизированном) режиме.

Одним из методов моделирования, позволяющих производить качественные прогнозы, является нейронная сеть [1; 2]. При построении нейросетевых моделей важно правильно выбрать архитектуру нейронной сети (типы и расположение нейронов и межнейронных связей) и настроить ее весовые коэффициенты. При выборе архитектуры нейронной сети необходимо найти «золотую середину». Недостаточная сложность ведет к большой, недопустимой ошибке, в то время как избыточная сложность ведет к восприятию шума в качестве сигнала, что лишает нейронную сеть (НС) свойства обобщения. Другим аспектом нейросетевой модели является ее обучение. Стандартные методы обучения могут не подходить для определенной архитектуры нейронной сети. Ввиду существования описанных выше проблем возникает задача создания программной системы, обеспечивающей автоматическое формирование и обучение нейронных сетей. Для этих целей могут быть использованы эволюционные алгоритмы, в частности, генетический алгоритм (ГА) и генетическое программирование (ГП) [3; 4].

Наиболее известным методом настройки весовых коэффициентов является метод обратного распространения ошибки, предложенный в 1974 году [5; 6]. Идея метода состоит в корректировке весовых коэффициентов межнейронных связей в направлении от выхода к входу. Градиент вычисляется согласно рекуррентным формулам. Важным замечанием для данного алгоритма является то, что он подходит только для полносвязных однонаправленных нейронных сетей. Для увеличения гибкости нейросетевого предиктора, а также сокращения времени обучения, необхо-

димо обеспечить возможность исключения некоторых межнейронных связей и появления связей между нейронами несоседних слоев. Для неполносвязных нейронных сетей неприменим алгоритм обратного распространения ошибки, однако для настройки нейронной сети может быть использован один из методов многомерной оптимизации, например, генетический алгоритм [3].

Самоконфигурируемый генетический алгоритм. ГА – оптимизационный метод, в основе которого лежит природная концепция эволюции, в частности, идея отбора наиболее пригодных индивидов, воспроизводства потомства и мутации. Генетические алгоритмы успешно применяются при решении задач поиска глобального оптимума, в том числе для сложных систем. Однако стандартный генетический алгоритм обладает таким недостатком, как зависимость качества решения от выбранных настроек (более ста комбинаций настроек даже для ГА, использующего только стандартные операторы) [7; 8].

Оптимальные значения параметров и настроек алгоритма изначально не известны, поэтому ставится задача отыскания набора таких параметров с целью увеличения производительности программной системы. Для решения такой задачи был предложен самоконфигурируемый генетический алгоритм SelfCGA [9], в котором применение комбинаций генетических операторов осуществляется исходя из успешности этих операторов на предыдущем поколении. Благодаря отсутствию необходимости хранить информацию о более ранних поколениях, при работе алгоритма не возникает проблемы больших затрат памяти, что было проблемой для более ранних вариантов ГА с автоматическим выбором настроек [10]. Принцип работы данного алгоритма можно представить в виде блок-схемы (рис. 1).

Вероятности применения операторов для порождения каждого следующего потомка не являются объектом эволюции и не включены в хромосому, что существенно сокращает размерность пространства оптимизации. Вычислительные ресурсы, требуемые для выполнения операторов, не увеличиваются с ростом числа поколений.

Описанный подход может быть применен для любых эволюционных алгоритмов, а также интегрирован с другими интеллектуальными информационными технологиями [11–13].

В целях проверки эффективности данного алгоритма проводилось тестирование стандартного генетического алгоритма и самоконфигурируемого генетического алгоритма SelfCGA, а также сравнение результатов. Показатели надежности (доли запусков, в которой алгоритм нашел решение с заданной точностью)

усреднялись по 20 прогонам на 12 тестовых функциях [14]. Графическое представление результатов дано на рис. 2, где сплошной линией иллюстрируется средняя надежность каждой из 120 комбинаций настроек,

пунктирная линия иллюстрирует среднюю надежность стандартного ГА (усреднение по 120 настройкам), а точечная линия иллюстрирует надежность SelfCGA.

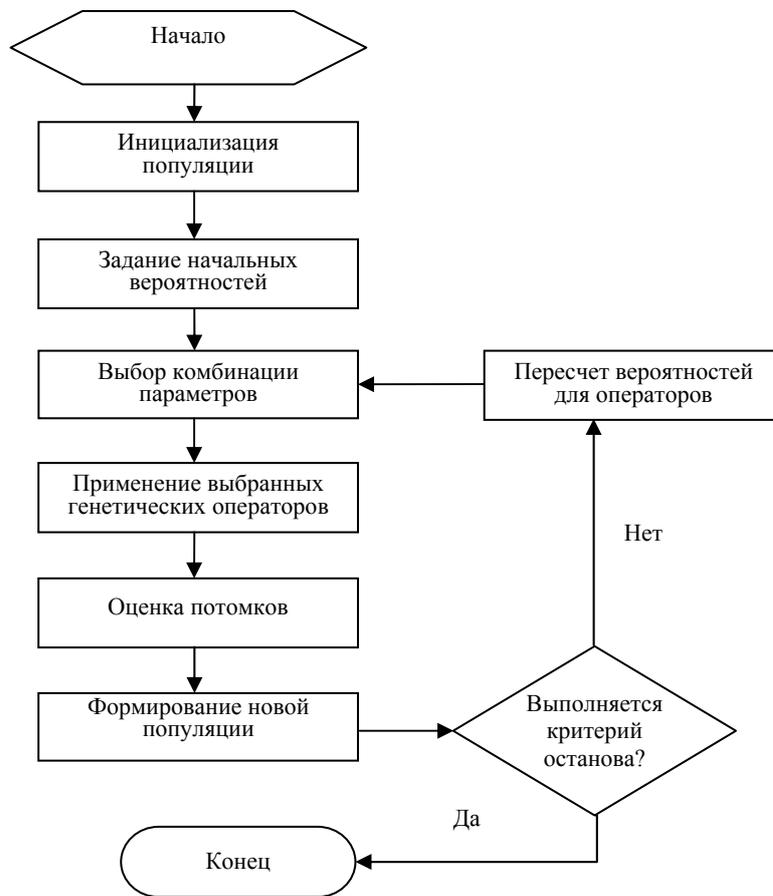


Рис. 1. Блок-схема самоконфигурируемого генетического алгоритма SelfCGA

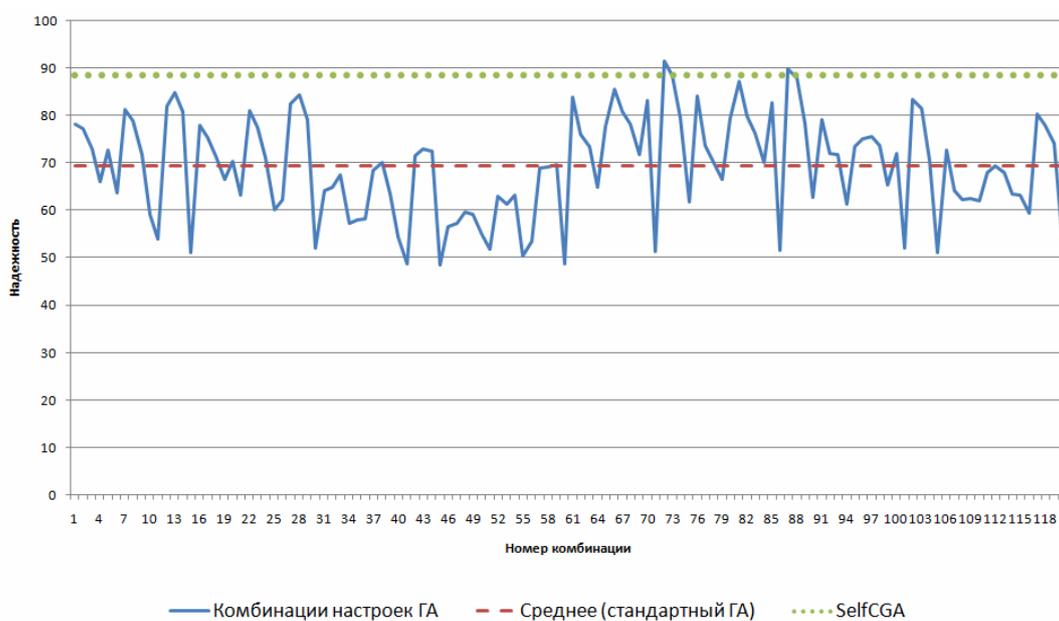


Рис. 2. Сравнение надежности генетических алгоритмов

Нетрудно видеть, что только две из ста двадцати комбинаций настроек дают надежность выше, чем у самоконфигурируемого алгоритма, причем это преимущество незначительно. Добавим, что при решении реальных задач эти наилучшие настройки заранее не известны, и пришлось бы проводить численные исследования, требующие чрезмерных затрат вычислительных ресурсов и времени. В этой связи можно утверждать, что на практике следует использовать SelfCGA в тех случаях, когда предполагается использование ГА.

Генетическое программирование. Как отмечалось ранее, качество решения задачи с помощью НС зависит от ее архитектуры. Метод проб и ошибок при выборе архитектуры НС в общем случае является неэффективным и требует не только значительных временных и вычислительных затрат, но и привлечения специалистов очень высокой квалификации. В связи с этим необходимо использовать процедуру автоматического формирования структур нейросетевых моделей. Для решения этой задачи в данной работе применяется алгоритм генетического программирования.

ГП является одним из видов эволюционных алгоритмов. ГП осуществляет поиск решений в пространстве деревьев. Основоположителем ГП считается Джон Коза, который внес значительный вклад в его развитие [4].

Для автоматического формирования нейронной сети с помощью алгоритма генетического программирования необходимо решить такие задачи, как кодирование нейронной сети в виде дерева и выбор функции пригодности для оценки решения. Для кодирования нейронной сети в виде дерева необходимо задать терминальное и функциональное множества таким образом, чтобы выполнялись условия замкнутости и достаточности.

Для создания нейронных сетей произвольной архитектуры необходимо применять способ кодирования деревьев, который позволяет создавать неполносвязные нейронные сети, сети с межслойными связями, а также нейроны с различными активационными функциями в пределах одного слоя.

Функциональное множество могут составлять две функции объединения нейронов:

1. Оператор, объединяющий ветви таким образом, что выходы нейронов левого поддерева становятся входом для нейронов правого поддерева.

2. Оператор, объединяющий нейроны левого и правого поддерева в один слой.

Терминальное множество могут составлять входы задачи (входные нейроны), а также нейроны с различными типами активационных функций (нейроны скрытых слоев).

Функция пригодности для дерева, кодирующего нейронную сеть, ставит в соответствие дереву число, обратно пропорциональное ошибке, сгенерированной по нему НС после заданного наперед количества итераций настройки весовых коэффициентов.

На рис. 3 представлена общая схема программной системы, разработанной и использованной в данной работе.

Решение задачи прогнозирования финансовых рядов. Для проверки эффективности разработанной программной системы была решена задача прогнозирования финансовых временных рядов. Следует подчеркнуть сложность такой задачи. Данные исследования основываются на техническом анализе ценовых движений, который подразумевает зависимость будущих изменений цен от прошлых. Однако существует гипотеза «случайных блужданий», предполагающая, что цены на рынке ценных бумаг меняются случайным образом и не могут быть предсказаны [15; 16].

Пусть дана информация о движении цен на акции транснациональной компании IBM за период с 2003 по 2014 гг. Данные предоставлены Yahoo Finance. Информация представлена в виде 15 переменных с временным промежутком в один торговый день (рабочие дни; время записи информации 00:00 EDT). Переменные могут быть разделены на 2 группы. Первая группа содержит фактические значения (цены открытия, закрытия, максимум, минимум, торговый оборот), вторая группа содержит значения технических индикаторов (RSI-2, EMA-7, EMA-50, EMA-200, MACD-26-12, Signal (MACD), DI+, DI-, DX, ADX), рассчитанных на основе значений первой группы.

Из общей выборки было выбрано 3 периода длиной в 7 месяцев каждый (5 для обучающей и 2 для тестовой выборки). Первый период характеризуется возрастающим движением цен на обучающей и тестовой выборках. Второй период характеризуется наличием возрастающего и убывающего движения цен на обеих выборках. Третий период характеризуется несовпадением трендов на обучающей и тестовой выборках. На данных выборках были произведены независимые запуски программной системы, которая была реализована в программе MATLAB 2013b [17].

Данная задача была также решена другими методами, такими, как многослойный перцептрон (MLP), полносвязная, однонаправленная нейронная сеть, алгоритм обучения – обратное распространение ошибки), радиально-базисная нейронная сеть (RBFN) и метод опорных векторов (SVMreg), который демонстрирует неплохие результаты в задачах прогнозирования финансовых рядов [18; 19].

Для получения результатов работы данных методов была использована система RapidMiner 5.3 [20]. Реализация метода опорных векторов представлена в расширении Weka (RapidMiner Weka Extension) [21].

В таблице показаны усредненные по 20 запускам результаты работы разработанной программной системы (GP+NN+SelfCGA) в сравнении с перечисленными выше методами.

На основании данных, представленных в таблице, можно сказать, что разработанная система успешно справляется с задачей прогнозирования финансовых временных рядов, показывая, в среднем, более высокие результаты по сравнению с конкурирующими подходами, представленными на рынке программных систем интеллектуального анализа данных. В квадратных скобках приведен разброс результатов в 20 прогонах. Предложенный подход проигрывает только методу опорных векторов и только на одной задаче, причем проигрыш не выглядит существенным.



Рис. 3. Схема программной системы

Сравнение результатов

Метод	Точность прогнозирования (выборка 1), %	Точность прогнозирования (выборка 2), %	Точность прогнозирования (выборка 3), %
MLP	61,36	61,36	52,27
SVMreg	63,63	65,90	59,09
RBFN	59,09	68,18	52,27
GP+NN+SelfCGA	68,21 [66,41 % – 70,11]	70,45 [69,51 % – 72,15]	56,81 [53,65 % – 59,43]

Заключение. Таким образом, разработанную программную систему можно считать одним из инструментов прогнозирования будущего состояния экономики, что позволяет более грамотно управлять финансами и снизить вероятность потери денежных средств за время разработки и внедрения инновационных проектов в ракетно-космической отрасли. Показано также, что применение эволюционных алгоритмов (ГА и ГП) делает нейронные сети одним из наиболее эффективных методов прогнозирования, так как теперь они не требуют вмешательства человека-эксперта в процесс их формирования и обучения.

Библиографические ссылки

1. Оссовский С. Нейронные сети для обработки информации. М. : Финансы и статистика, 2002. 344 с.
2. Anderson D., McNeill G. Artificial neural networks technology // DACS report. 1992. 87 p.
3. Haupt R. L., Haupt S. E. Practical Genetic Algorithms. 2 ed. Wiley, 2004. P. 261.
4. Koza J. R. Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, 1992.
5. Галушкин А. И. Синтез многослойных систем распознавания образов. М. : Энергия, 1974.
6. Werbos P. J., Beyond regression: New tools for prediction and analysis in the behavioral sciences. Ph. D. thesis. Cambridge : Harvard University, 1974.
7. Wright A. Genetic algorithms for real parameter optimization. Foundations of Genetic Algorithms. San Mateo, Morgan Kaufmann, 1991. Pp. 205–218.
8. Eiben A. E., Smith J. E., Introduction to evolutionary computing. Berlin : Springer, 2003.

9. Semenkin E., Semenkina M. Self-configuring genetic algorithm with modified uniform crossover operator // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 7331 LNCS (PART 1). 2012. Pp. 414–421.
10. Gomez J. Self Adaptation of Operator Rates Evolutionary Algorithms // Genetic and Evolutionary Computation – GECCO 2004 : Lecture Notes in Computer Science. Vol. 3102. 2004. Pp. 1162–1173.
11. Semenkina M., Semenkin E. Hybrid Self-configuring Evolutionary Algorithm for Automated Design of Fuzzy Classifier // Advances in Swarm Intelligence / Y. Tan [et al.] (Eds.). ICSI 2014. Part 1. LNCS 8794. 2014. P. 310–317.
12. Burakov S. V., Semenkin E. S. Solving variational and Cauchy problems with self-configuring genetic programming algorithms // International Journal of Innovative Computing and Applications. 2013. Vol. 5, iss. 3. P. 152–162.
13. Semenkin E., Semenkina M. Classifier Ensembles Integration with Self-Configuring Genetic Programming Algorithm // Adaptive and Natural Computing Algorithms : Lecture Notes in Computer Science. Vol. 7824. Berlin, Heidelberg : Springer-Verlag, 2013. P. 60–69.
14. Molga M., Smutnicki Cz. Test functions for optimization need. 3 kwietnia. 2005.
15. Fama E. F. Random walks in stock market prices // Financial Analysts Journal. 1965. 21(5). P. 55–59.
16. Cootner P. H. The random character of stock market prices. MIT Press. 1964.
17. R2013b – Updates to the MATLAB and Simulink product families [Электронный ресурс]. URL:

http://www.mathworks.com/products/new_products/release2013b.html (дата обращения: 18.01.2015).

18. Zhang Y. Prediction of Financial Time Series with Hidden Markov Models. Simon Fraser University, 2004.

19. Fedotov D. V. On financial time series prediction using neural network models and evolutionary algorithms // Materials of the Third International Workshop on Mathematical Models and its Applications (IWMMMA-2014). 2014. P. 3.

20. Predictive Analytics, Data Mining, Self-service, Open source – RapidMiner [Электронный ресурс]. URL: <http://rapidminer.com/> (дата обращения: 20.10.2014).

21. Rapid-I Market Place Weka Extension for RapidMiner [Электронный ресурс]. URL: http://marketplace.rapid-i.com/UpdateServer/faces/product_details.xhtml?productId=rmx_weka (дата обращения: 20.10.2014).

References

1. Ossovskiy S. *Neyronnye seti dlya obrabotki informatsii* [Neural networks for information processing]. Moscow, Finansy i statistika Publ., 2002, 344 p.

2. Anderson D., McNeill G. Artificial neural networks technology. *DACS report*. 1992, 87 p.

3. Haupt R. L., Haupt S. E. *Practical Genetic Algorithms*, 2ed., Wiley, 2004. P. 261.

4. Koza J. R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, 1992.

5. Galushkin A. I. *Sintez mnogosloynnykh sistem raspoznavaniya obrazov* [Synthesis of the multilayer pattern recognition systems]. Moscow, Energiya Publ., 1974.

6. Werbos P. J., Beyond regression: New tools for prediction and analysis in the behavioral sciences. *Ph. D. thesis*, Harvard University, Cambridge, MA, 1974.

7. Wright A. Genetic algorithms for real parameter optimization. *Foundations of Genetic Algorithms*. San Mateo, CA: Morgan Kaufmann, 1991, p. 205–218.

8. Eiben A. E., Smith J. E., Introduction to evolutionary computing. Springer, 2003, Berlin, Germany.

9. Semenkin E., Semenkina M. Self-configuring genetic algorithm with modified uniform crossover operator, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7331 LNCS (PART 1), 2012, p. 414–421.

10. Gomez J. Self Adaptation of Operator Rates Evolutionary Algorithms, *Genetic and Evolutionary Computation, GECCO 2004. Lecture Notes in Computer Science*. 2004, Vol. 3102, p. 1162–1173.

11. Semenkina M., Semenkin E. Hybrid Self-configuring Evolutionary Algorithm for Automated Design of Fuzzy Classifier, *Advances in Swarm Intelligence*. Y. Tan et al. (Eds.): ICSI 2014, Part 1, LNCS 8794, 2014, p. 310–317.

12. Burakov S. V., Semenkin E. S. Solving variational and Cauchy problems with self-configuring genetic programming algorithms, *International Journal of Innovative Computing and Applications*. 2013, Vol. 5, Iss. 3, p. 152–162.

13. Semenkin E., Semenkina M. Classifier Ensembles Integration with Self-Configuring Genetic Programming Algorithm, *Adaptive and Natural Computing Algorithms. Lecture Notes in Computer Science*, 2013, Vol. 7824. Springer-Verlag, Berlin, Heidelberg, p. 60–69.

14. Molga M., Smutnicki Cz., 2005. Test functions for optimization need. 3 kwietnia.

15. Fama E. F. Random walks in stock market prices, *Financial Analysts Journal*. 1965, no. 21(5), p. 55–59.

16. Cootner Paul H. *The random character of stock market prices*. MIT Press. 1964.

17. R2013b – Updates to the MATLAB and Simulink product families, Available at: http://www.mathworks.com/products/new_products/release2013b.html (accessed 18.01.2015).

18. Zhang Y. *Prediction of Financial Time Series with Hidden Markov Models*, Simon Fraser University, 2004.

19. Fedotov D. V. On financial time series prediction using neural network models and evolutionary algorithms // Materials of the Third International Workshop on Mathematical Models and its Applications (IWMMMA-2014), 2014, P. 3.

20. Predictive Analytics, Data Mining, Self-service, Open source – RapidMiner. Available at: <http://rapidminer.com/> (accessed 20.10.2014).

21. Rapid-I Market Place Weka Extension for RapidMiner, Available at: http://marketplace.rapid-i.com/UpdateServer/faces/product_details.xhtml?productId=rmx_weka (accessed 20.10.2014).

© Федотов Д. В., Семенкин Е. С., 2014