Proceedings of the International Multiconference on
Computer Science and Information Technology, pp. 81 – 88

ISBN 978-83-60810-22-4
ISSN 1896-7094

Trading rule discovery on Warsaw Stock Exchange using coevolutionary algorithms

Paweł B. Myszkowski
Wrocław University of Technology,
Wyb.Wyspiańskiego 27, 51-370 Wrocław. Poland
Email:pawel.myszkowski@pwr.wroc.pl

Łukasz Rachwalski
Wrocław University of Technology,
Wyb.Wyspiańskiego 27, 51-370 Wrocław. Poland
Email: 166499@student.pwr.wroc.pl

Abstract — **This paper presents an application of coevolutionary algorithms to rule discovery on stock market. We used genetic programming techniques with coevolution in financial data mining process. There were tested a various approaches to include coevolution aspects in task of build trading rule (buy and sell decision). Trading rules are based on technical and fundamental indicators included in decision tree and were tested on Warsaw Stock Exchange historical data.**

I. Introduction

IN RECENT years, a number of artificial intelligence approaches have been suggested for applications in financial data mining tasks, especially stock market analysis. Predicting share prices, share value estimation or trading rule generation are the most interesting tasks that are being solving not only by neural networks techniques, genetic programming or evolutionary algorithms but much more.

A. Related works

In [2], [7] are proposed a hybrid approach based on evolutionary algorithm and artificial neural networks for predicting trends of stock market indicators. Work [9] presents intelligent decision support system for stock market investor. System is based on genetic algorithms in cooperation with artificial neural networks and fuzzy logic and was successfully tested on Taiwan stock market.

An another interesting approach described in [3] connects artificial neural networks and case-based reasoning. Its working schema can be shortly presented as follows: (a)monitoring of potential interesting shares (b)artificial neural networks decide about time of buy/sell transaction (c)verification of observed results in comparison to artificial neural networks prediction using historical cases.

A related work [16] shows swarm intelligence application for generating artificial neural networks for supporting investing decision. An artificial neural networks are applied for daily quotation analysis and buy/sell signal generation.

In paper [10] is proposed a machine learning system called TPP (Turning Points Prediction) as framework based on chaotic dynamic analysis and neural network modeling for prediction peaks and troughs of indexes. Such tool is able to help investor in market trend recognition and profitable opportunities discovering. Presented experimental results showed that the developed solution can be helpful to make profitable transactions.

In paper [15] is presented induction learning system based on evolutionary algorithms applied to profitable rule discovery from Warsaw Stock Exchange [1] (WSE) historical data. The obtained rule has form of decision tree and contains fundamental and technical analysis indicators as well. Each individual consists of two decision trees: buy and sell tree, and a given decision tree defines condition(s) that should be satisfied to generate buy/sell signal for selected share. Defined fitness function is based on return of profit generated by investing strategy.

In this work approach is more complex than presented in [15]. The main difference is coevolution usage, where trading rules were splitted into two independent populations: one consists of buy rules, the other one includes selling rules. Our prerequisites are that some works (e.g. [5], [12], [13]) shows that evolutionary algorithms with coevolution mechanism gives higher efficiency and more suitable solutions. Evolution of each population works independent, but there are one connecting aspects (coevolution): individual fitness function of one population is connected to individual of second population – in order to evaluate selling rule strongly is needed buy rule work (and vice versa). The fitness function is defined as return of initial capital and its corresponding to profit that rules have gained in its transactions.

Our main goal is defined as follows: to experimentally test evolutionary algorithm efficiency in financial data mining task and comparing results with [15] including coevolution aspects. The second goal is to experimentally testing if obtained trading rules have knowledge independent on training data (it there is any generalization) and if there is any influence to training data stock market trend. For instance if rule discovered in uptrend duration can give profit on another period in downtrend or horizontal trend, where explicit trend does not exist. An additional goal is to investigate efficiency improvement using coevolutionary algorithms.

The paper is organized as follows: the next section describes proposed evolutionary algorithm approach. There are defined an individual representation, genetic operators and fitness function method. Also evolutionary algorithms

[1] *WSE* (*Warsaw Stock Exchange*) is the largest stock exchange in eastern Europe, located in Warsaw Poland and opened on April 16, 1991 (internet source: http://www.gpw.pl/)

extensions are presented, especially coevolution. Section III describes done experiments (long and short time periods, also various type of trends) and observed results. Conclusions and future research directions are given in the last section.

II. Metodology

A task is defined as follows: evolutionary algorithm in the base of historical data from WSE builds strategy based on two types of rule: buy and sell rule. Given rule defines condition(s) that should be satisfied by company share to generate buy/sell signal. An important issue is obtained knowledge generalization of in financial data mining task. Thus there are used data separation: training and testing data [4] , where rules generated on training data are experimentally tested on testing data. Also into consideration was taken character of data and its influence into possible rule earnings.

We investigated evolutionary algorithm in three approaches: typical evolutionary algorithm, evolutionary algorithm with fitness function evaluation and coevolutionary algorithm. We investigated also fundamental analysis including.

A. Evolutionary Algorithm (EA)

To apply Evolutionary Algorithm metaheuristics we have to define individual representation (to decide about problem representation method), genetic operators (decision about problem elements manipulators) and fitness function form to evaluate solution propositions.

```
i:=0;
initialise(pop i );
evaluate(pop i );
    while (!stop_condition)
        pop i+1 := selection(pop i );
        pop i+1 := crossover(pop i+1 );
        pop i+1 := mutation(pop i+1 );
        evaluate(pop i+1 );
        i:=i+1;
```

EA starts with initial population (usually created randomly). Next, the individuals in population are evaluated – each individual receives fitness function value that corresponds to quality of its proposition of given problem solution ([1], [6], [11]). Next step checks if stop conditions are not met: usually it is limit of generations and the best individual fitness value is acceptable (success). If stop criteria is not met EA runs selection procedure that defines a seed for new generation and provides communication between individuals (by crossover operator) and the independent search by mutation operator. The whole process is repeated until any stopping condition is met.

1) Individual representation

Each individual contains a portfolio of its owned shares and two independent decision trees [8]: the first one decides about purchasing shares (BUY), the other one about selling (SELL). The individual operates on simulated stock market for specific amount of sessions and decides about its transactions. During each session it can buy or sell shares of any company, where amount of money is divided on *N* equal packets. In a given moment the individual can have only one packet of given company shares.

A decision tree consists of two types of nodes: logical one and terminal. A logical node contains logical operator (OR / AND) which joins terminal nodes and indeed generates logical expression. Each logical node operates on two sub trees, returns a logical value (true or false) and can be nested.

A terminal node consists of three elements: (1) indicator type (technical or fundamental), (2) comparison operator (less-than "<" or more-than „>") and (3) floating-point number. The whole decision tree is a logical expression, which tells if a given share conditions are satisfied and can be bought/sold.

The example of such trading rule (represented as a decision tree and logical expression) is presented on Fig 1. Presented decision tree looks for share that RSI indicator is higher than 80,5, and also satisfies one of conditions: ROC is lower than 0,97 or DMA is lower than -7,29.

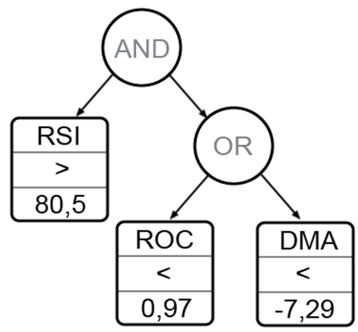

Fig 1: Example of decision tree as trading rule. Its textual representation as logical expression can be presented as follows:

RSI>80,5 AND (ROC<0,97 OR DMA<-7,29)

Such representation is intuitive, very useful in operating by evolutionary algorithm and simple in analyzing by potential human investor.

2) Selection and initialization method

We used a proportional selection (so-called roulette wheel) that prefers in population individuals with fitness function is higher that average. An elite parameter was also tested in our approach.

The initial population was created by random method but limited by decision tree size; we used a maximal tree size parameter to avoid its overgrowing.

3) Genetic operators

We used genetic programing type of one-point crossover that connects randomly corresponding (buy/sell) decision trees. Mutation operators have been split into three types: node modification (*NM*), new node insertion (*NI*) and node deletion (*ND*). Each of them is connected to usage probability value and works on randomly selected node of decision tree.

The *NM* operator works on values stored in node (an indicator or a logic operator, a comparison operator and a value). The higher change of probability has (descending): a node value (according to normal distribution), comparison operator and type of indicator (or logical operator).

The *NI* operator indeed inserts two nodes (a terminal and logical operator) to keep a given tree coherent. Selected node of tree become an the left descendant of newly inserted logical node, and the second inserted nodes (terminal one) becomes the right descendant to keep decision tree coherent.

The *ND* operator removes a randomly selected node including a set of its all descendants (indeed the whole sub tree) to keep tree coherent in its place is inserted a new randomly generated terminal node.

4) Fitness function

In our approach, evaluation of each population of individuals requires stock market simulation in given historical time period. For each stock market session and each quotation we need to ask an individual for sell/buy decision. The individual decides independently about transaction: to buy or sell owned share and it causes a loss or yields a profit. To evaluate each individual it owns initial capital $CASH_{start}$ and each financial decision has influence on final capital return $CASH_{stop}$ calculated at the end of time period. If an individual owns any shares at the end of time period they are sold. Total value of its fitness function is generated by percentage return of initial capital:

$$fitness = \frac{CASH_{stop}}{CASH_{start}} \tag{1}$$

Because of decision tree application and EA tendency to overgrown trees we use tree depth limit and each node more than limit causes its fitness function value reduction by 2% as penalty (this value was set experimentally). Another reduction of fitness function value is connected with a brokerage and equals 3% of each buy/sell operation value (this value is set experimentally).

B. Extended Evolutionary Algorithm (ExEA)

We extended standard fitness evaluation function by building pairs of the best buy-sell individual. This process is based on tournament of given percentage (as *TreeMatching* parameter) individuals in population. For example, *TreeMatching* equals to 10% means that given individual's decision buy tree are evaluated with selling decision trees of 10% randomly chosen individuals of given population. The best selling rule is inserted into given individual.

C. Coevolutionary Algorithm (CoEA)

As Coevolutionary Algorithm is based on cooperation of two (or more) populations, in our approach we have populations: (1) buy and (2) sell decision trees. The individual of one population is connected to individual of second population only by evaluation function value to make best profit: to buy "interesting" share in "good" moment, and sell with "good" profit.

To evaluate individuals of two coevolving population we link them into pairs buy-sell and profit of such strategy is calculated. The final value of fitness function given individual is average profit obtained by given rule connected to *treeMatching* (%) individuals of convolving population (for each pair the fitness function value is evaluated separately).

III. Experiments and results

We investigated many experiments on WSE historical data to verify efficiency of implemented approaches. We used data [2] from different character time periods, but only quoted companies in whole analyzed time period (debutants were omitted). Our research methodology assume that EA firstly learns to discover trading rules using data from training period, and then its efficiency is tested on selected testing period [4] .

All tests have been carried for different version of algorithm many times using different parameters' values. A following approaches were tested:

- evolutionary algorithm (*EA*),
- evolutionary algorithm using financial analysis indicators (*EA + fund*),
- extended evolutionary algorithm with parameter *treeMatching* = 10% (*ExEA* (10%)),
- coevolutionary algorithm CoEA, *treeMatching* = 10% (*CoEA* (10%)),
- coevolutionary algorithm, *treeMatching* = 30% (*CoEA* (30%))
- coevolutionary algorithm with *treeMatching* = 10% using financial analysis indicators (*CoEA* (10%) + *fund* .)

Results of presented approaches EA and EA+fund can be analyzed as substitute of methodology defined in [15]. However results presented there are based on other set of indicators (a fundamental and a technical) also were tested on different periods so its results cannot be strictly compared.

Presented approaches builds decision trees taking into consideration following technical indicators: closing price (*INX*), volume (*VOL*), *ROC(5)*, *ROC(10)*, *RSI(5)*, *RSI(10)*, *DMA(5, 20)* and *MACDO(12, 26, 9)* [14]. A value in brackets defines a number of sessions used to compute a given indicator value. Some of presented approaches takes into consideration also financial analysis indicators (calculated in base of company quarter reports) as follows: operating profit margin (so called return on sales), gross profit margin and net profit margin [17].

For investigations, we selected nine short test periods – three for each trend on stock market: uptrend (denoted as *Ux*), downtrend (denoted as *Dx*) and horizontal trend (denoted as *Hx*). We also selected three training periods, one for each type of trend. Training periods are additionally denoted by zero sign, e.g. *U0* is training uptrend period. Testing periods are enumerated by integers number from 1 to 3. All selected periods are listed below with an additional information: symbol, length (number of stock market sessions), number of quoted companies, change of main WSE index WIG during given period to show its character:

```
U0: 25 quoat., 188 shares, WIG  +8,90%
H0: 21 quoat., 253 shares, WIG  +0,68%
D0: 28 quoat.,  91 shares, WIG  -8,92%
U1: 21 quoat., 197 shares, WIG  +9,80%
```

[2]Source: http://www.bossa.pl

Table I. Test results of individuals trained on U0 period

algorithm type	U1 (+9,8%)						H1 (+0,95%)						D1 (-7,28%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	17,98	6,72	13,33	1,14	**10,23**	6,74	3	3,93	6,1	1,03	**11,85**	6,34	-3,88	2,83	-9,69	0,53	**-9,19**	1,88
EA + fund	14,63	5,7	19,44	2,02	**22,04**	5,01	3,66	3,65	3,32	1,1	**16,35**	4,32	-8,8	8,93	-14,08	0,95	**-16,8**	2,41
ExEA (10%)	12,22	0	16,08	3,5	**16,48**	10,6	2,24	5,01	7,01	1,02	**10,39**	5,58	-11,42	3,47	-11,21	0,68	**-12,41**	2,82
CoEA (10%)	18,34	13,22	13,54	1,57	**11,28**	7,05	10,09	8,84	7,02	1,03	**12,22**	5,56	-11,02	6,63	-10,08	0,58	**-9,13**	1,65
CoEA (30%)	21,14	13,94	12,92	1,67	**13,17**	4,63	8,69	8,57	7,09	1,59	**15,18**	6,62	-10,31	5,98	-10,02	0,59	**-9,48**	1,23
CoEA (10%) + fund	10,75	5,68	22,25	2,17	**23,28**	4,86	-0,4	3,36	4,66	1,53	**17,23**	4,4	-9,11	2,18	-14,79	0,69	**-17,73**	1,85
algorithm type	U2 (+12,83%)						H2 (+0,3%)						D2 (-7,49%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	5,26	3,91	2,57	0,7	**2,99**	2,79	-1,25	4,46	-1,37	0,68	**-1,35**	2	-8,28	2,84	-10,25	0,52	**-11,54**	1,91
EA + fund	4,78	4,57	2,92	1,13	**11,65**	5,37	1,13	2,71	0,6	0,38	**1,79**	1,97	-6,69	3,82	-5,43	1,03	**-10,16**	2,07
ExEA (10%)	0,82	2,72	3,47	1,93	**3,71**	4,09	0	2,29	-0,74	1,11	**-3,33**	3,28	-8,22	3,77	-10,3	1,48	**-12**	3,04
CoEA (10%)	1,38	2,54	2,63	0,84	**2,21**	2,59	1,57	6,47	-1,5	0	**-1,38**	2,79	-13,39	4,31	-10,38	0,68	**-11,58**	1,83
CoEA (30%)	1,22	3,34	2,56	0,8	**1,12**	1,57	-1,85	3,77	-1,16	1,07	**-1,14**	3,54	-11,24	4,24	-10,48	0,83	**-13,51**	1,74
CoEA (10%) + fund	0,98	1,72	4,02	1,5	**13,28**	3,22	0,99	1,26	0,26	0,58	**-0,39**	3,61	-9,06	2,53	-6,19	0,9	**-11,24**	2,34
algorithm type	U3 (+11,95%)						H3 (+0,02%)						D3 (-5,03%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	8,65	4,53	8,76	1,03	**10,2**	3,21	3,7	2,27	-0,12	0,6	**0,39**	2,36	0,23	2,69	-1,54	0,72	**-4,33**	1,95
EA + fund	5,09	3,92	4,3	0,61	**4,96**	2,28	1,27	2,04	-3,36	0,76	**-5,5**	2,45	-1,34	3,67	-6,9	1,17	**-11,44**	2,6
ExEA (10%)	6,15	3,96	6,82	1,57	**8,96**	4,01	0,18	2,01	-1,43	1,08	**-2,87**	4,01	0,07	2,19	-1,08	0,66	**-3,47**	2,96
CoEA (10%)	8,05	5,24	8,61	0,99	**9,99**	3,55	-3,66	3,17	-0,41	0,66	**-0,4**	3,39	-2,58	2,11	-1,76	0,95	**-4,02**	2,64
CoEA (30%)	7,82	6,2	8,26	0,93	**9,52**	2,93	-3,55	3,99	-0,36	0,68	**-0,9**	2,78	-1,37	1,8	-1,84	0,66	**-4,45**	2,55
CoEA (10%) + fund	5,04	2,8	4,72	0,58	**3,65**	2,25	0,23	1,4	-3,84	0,58	**-6,44**	2,18	0,13	1,27	-7,21	1,04	**-11,93**	2,25

* In the braces is given the *treeMatching* parameter value.

Table III. Test results of individuals trained on H0 period

algorithm type	U1 (+9,8%)						H1 (+0,95%)						D1 (-7,28%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	15,93	6,21	8,17	1,6	**17,99**	12,24	3,95	3,36	2,46	0,79	**5,81**	3,76	-5,83	6,72	-6,06	1,28	**-8,76**	5,08
EA + fund	11,43	4,83	9,24	1,84	**18,51**	10,95	2,65	2,62	1,93	0,78	**6,66**	4,17	-4,9	5,63	-8,02	1,45	**-11,45**	5,48
ExEA (10%)	8,78	7,1	9,71	1,54	**11,92**	7,09	1,52	3,76	2,52	0,85	**6,68**	3,96	-9,06	5,58	-8,11	1,01	**-6,53**	3,54
CoEA (10%)	0	6,15	8,89	2	**22,66**	11,56	7,12	3,25	2,73	0,78	**9,17**	5,32	-14,96	3,88	-7,16	1,39	**-13,8**	3,79
CoEA (30%)	19,48	7,01	8,73	1,26	**15,72**	12,24	7,64	4,2	2,55	0,87	**6,11**	4,5	-13,96	5,21	-7,24	1,28	**-7,4**	3,99
CoEA (10%) + fund	2,93	8,01	9,58	1,55	**18,26**	12,15	7,33	4,4	2,11	0,75	**5,1**	4,96	-16,58	5,05	-8,07	1,11	**-11,51**	4,28
algorithm type	U2 (+12,83%)						H2 (+0,3%)						D2 (-7,49%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	6,69	3,95	3,26	0,91	**7,52**	4,5	0,59	3,66	0,65	0,83	**-1,54**	2,99	-6,38	2,18	-5,24	1,3	**-6,51**	3,99
EA + fund	5,37	4,01	2,73	0,83	**4,14**	5,95	-0,72	2,63	0,69	0,85	**-1,79**	3,34	-7,11	2,69	-4,82	1,11	**-8,21**	4,8
ExEA (10%)	3,23	4,22	3,86	1,16	**5,7**	4,61	0,8	2,57	0,28	0,29	**1,36**	1,93	-6,62	4,17	-6,34	1,08	**-5,22**	4,42
CoEA (10%)	8,08	5,23	3,33	1,19	**4,01**	7,93	-1,52	3,79	0,55	0,74	**-1,74**	3,55	-12,38	7,61	-6,16	1,09	**-11,98**	3,94
CoEA (30%)	6,8	6,13	3,71	1,47	**6,94**	6,23	-2,5	3,8	0,44	0,53	**-0,57**	2,74	-11,53	5,95	-6,17	1,16	**-6,62**	5,05
CoEA (10%) + fund	7,94	5,18	3,01	0,97	**7,49**	7,42	-1,89	4,17	0,32	0,45	**-2,27**	3,27	-10,95	6,86	-5,13	1,06	**-8,84**	5,18
algorithm type	U3 (+11,95%)						H3 (+0,02%)						D3 (-5,03%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	6,64	2,86	3,26	1,02	**2,6**	2,3	1,58	2,63	0,42	0,24	**0,44**	1,4	0,76	3,27	0,19	0,36	**-0,14**	1,84
EA + fund	5,97	3,83	3,42	0,75	**5,67**	4,38	2,32	1,38	-0,21	0,42	**-0,12**	2,99	0,04	1,48	-1,2	0,76	**-1,86**	2,75
ExEA (10%)	3,41	2,81	3,74	0,61	**2,99**	3,27	0,9	1,26	0,55	0,31	**0,95**	0,97	0,23	1,32	0,21	0,46	**-0,42**	1,19
CoEA (10%)	6,04	4,38	3,61	0,74	**7,91**	5,05	-2,75	2,67	-0,04	0,28	**-0,76**	3,53	-1,71	1,84	-0,07	0,39	**-0,65**	3,21
CoEA (30%)	5,46	3,29	3,51	0,83	**4,15**	4,46	-0,2	2,79	-0,06	0,29	**-1,07**	2,29	1,9	2,71	0,04	0,42	**-1,08**	1,67
CoEA (10%) + fund	5,93	3,85	3,7	0,9	**5,06**	3,84	-1,85	3,15	-0,33	0,44	**-1,06**	2,54	-3,04	2,81	-1,04	0,58	**-2,01**	3,02

Table II. Test results of individuals trained on D0 period

algorithm type	U1 (+9,8%)						H1 (+0,95%)						D1 (-7,28%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	11,45	3,19	9,91	1,24	**11,4**	5,59	2,52	2,68	2,06	0,73	**3,37**	3,07	-1,16	2,87	-7,66	0,98	**-9,79**	1,6
EA + fund	12,24	3,48	9,81	1,03	**8,4**	4,7	1,54	1,94	1,7	0,69	**3,35**	3,8	-2,45	4,54	-8,58	1,04	**-10,83**	2,23
ExEA (10%)	7,52	5,62	7,53	0,82	**8,29**	3,92	2,98	4,88	1,18	0,49	**3,79**	3,04	-10,1	4,45	-9,24	0,67	**-11,2**	2,35
CoEA (10%)	17,07	8,95	10,68	1,55	**10,35**	5,23	6,62	4,39	2,63	0,96	**1,42**	3,62	-16,08	0,29	-8,42	0,91	**-11,26**	2,75
CoEA (30%)	23,5	10,3	10,2	1,47	**7,56**	4,5	8,68	4,36	2,13	0,72	**0,89**	3,21	-17,55	3,74	-8,38	0,93	**-9,72**	2,03
CoEA (10%) + fund	27,34	9,07	9,96	1,32	**12,61**	6,21	5,17	5,22	1,84	0,82	**4,98**	4,6	-19,11	3,11	-8,76	1,1	**-11,93**	3,76
algorithm type	U2 (+12,83%)						H2 (+0,3%)						D2 (-7,49%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	5,31	2,42	3,73	0,9	**1,82**	2,97	0,09	2,16	0,81	0,45	**1,07**	2,14	-8,29	1,65	-6,84	0,95	**-9,69**	2,69
EA + fund	4,94	2,69	3,28	0,74	**0,89**	2,34	-0,26	1,56	0,62	0,49	**0,6**	1,79	-8,17	2,34	-6,51	0,96	**-9,26**	2,58
ExEA (10%)	2,45	5,13	1,39	0,8	**0,33**	2,59	0,64	1,82	1,07	0,34	**1,83**	1,3	-8,69	4,53	-8,45	0,76	**-10,86**	2,31
CoEA (10%)	3,52	4,78	4,12	0,95	**0,36**	2,05	-1,93	2,73	0,78	0,51	**1,08**	1,63	-13,27	3,79	-7,23	0,78	**-11,59**	2,79
CoEA (30%)	8,49	6,87	3,79	0,87	**1,31**	3,11	-2,54	3,22	0,69	0,43	**1,12**	2,17	-12,58	3,69	-7,09	0,96	**-10,48**	1,47
CoEA (10%) + fund	2,96	5,22	2,92	0,98	**0,85**	2,85	0,33	2,59	0,75	0,47	**1,93**	2,71	-6,64	5,38	-6,22	0,77	**-10,5**	3,02
algorithm type	U3 (+11,95%)						H3 (+0,02%)						D3 (-5,03%)					
	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev	min	dev	avg	dev	max	dev
EA	8,67	2,13	3,89	0,62	**8,45**	4,61	3,32	1,56	0,31	0,29	**1,51**	1,56	1,11	1,31	0,15	0,25	**1,69**	2,04
EA + fund	7,05	3,31	3,84	0,62	**7,54**	5,64	1,74	1,44	0,09	0,35	**1,47**	1,69	0,86	2,72	-0,18	0,47	**1,81**	2,79
ExEA (10%)	5,17	4,03	5,1	0,56	**6,24**	4,24	0,69	1,91	0,6	0,32	**1,92**	1,83	-0,82	1,81	-0,07	0,29	**2,87**	2,35
CoEA (10%)	10,66	6,28	3,45	0,52	**4,81**	3,37	0,31	2,89	-0,05	0,23	**1,82**	1,94	-2,44	4,2	0,21	0,25	**2,28**	2,17
CoEA (30%)	5,67	4,69	3,3	0,68	**5,22**	4,33	0,75	2,48	0,12	0,28	**1,31**	1,89	-2,81	3,44	0,15	0,26	**1,1**	1,58
CoEA (10%) + fund	4,66	2,67	3,38	0,68	**4,62**	3,09	-3,55	3,58	-0,2	0,33	**0,53**	1,63	-9,77	4,83	-0,3	0,47	**2,03**	2,31

H1: 29 quoat., 198 shares, WIG +0,95%
D1: 27 quoat., 205 shares, WIG -7,28%
U2: 38 quoat., 254 shares, WIG +12,83%
H2: 27 quoat., 270 shares, WIG +0,30%
D2: 27 quoat., 264 shares, WIG -7,49%
R3: 34 quoat., 82 shares, WIG +11,95%
B3: 28 quoat., 90 shares, WIG +0,02%
S3: 25 quoat., 83 shares, WIG -5,03%

To evaluate quality of obtained trading rule we compared its profit with main stock index *WIG* change. The value of *WIG* represents the simplest strategy called „buy-and-hold”, which is based on buying shares at the beginning of the given period and selling them out at the end of the period. Result of this strategy is a reference in evaluating the others strategies efficiency.

A. *Tests*

Each of the developed EA approach was running 30 times for all of three training periods. The individuals of last generation population were additionally tested on nine test periods. The averaged results of 30 test runs of developed algorithms are given in Table I-II. For each test period the lowest (*min*), the average (*avg*) and the highest (*max*) profit ratio in population is given. Next to the profit ratio standard deviation (*dev*) is given.

The population size was set to 100 individuals which evolves for 200 generations (experimentally set as a stop condition – the greater value causes occurance of overfitting and overtraining phenomenons). We use the roulette selection and the elite parameter equals to 1 which means that only the best individual survives without changes to the next generation. Values of particular parameters were determined experimentally: crossover probability Px =0,6, mutation probability Pm =0,4 and within mutation: node modification Pm_{NM} =0,65, new node insertion Pm_{NI} = 0,15 and the node deletion Pm_{ND} =0,2. Each individual received 10 thousands of virtual polish zlotych (*PLN*) as a start cap ital and could own at most 10 (*N* parameter) packets of shares at a given moment.

In all training periods *ExEA* reaches the highest fitness value (significantly outperforming other approaches), however results from test periods were much more diversified (depending on period) so it is hard to draw any far going conclusions about efficiency of developed approaches. With the aid of data presented on Table I - Table III we cannot point the best unrivaled approach. We can draw a conclusions as follows:

- *EA* had the lowest losses in downtrend periods,
- *CoEA(10%)* trained on *H0* achieves distinctive profit only in periods *U1* and *U3* .
- approaches indicating financial indicators (*EA+fund* , *CoEA+fund*) achieves the highest (or close to highest) profits in periods up- and horizontal trend periods *(U1* , *H1* , *U2* and *H2),* but it performed poorly in downtrend periods.

The last conclusion we find very interesting: we suppose that individual using financial analysis buys shares while downtrend period because of very cheap shares (according to the "value investing" theory). The set of individuals with the highest profits gained in short period tests we selected for another tests: long time tests.

We selected two individuals of 30 runnings of each developed approach of each training period (total 36 best individuals were selected). These individuals profits reached in 9 short test periods are presented in Table IV: where the last column a total profit of selected individual is given. In all these 9 periods „buy-and-hold" strategy would yield a profit

Table IV. Short time test results of best obtained individuals

Training Period	Algorithm type	No.	U1 (9,85%)	H1 (0,95%)	D1 (-7,28%)	U2 (12,83%)	H2 (0,30%)	D2 (-7,49%)	U3 (12,00%)	H3 (0,02%)	D3 (-5,03%)	Total Profit [%]
U0	EA	1	44,16	12,37	-10,47	0,66	-4,52	-8,54	14,84	-0,46	-5,95	42,09
		2	11,29	35,4	-6,49	0,98	0,11	-12,65	12,78	2,03	-5,58	37,87
	EA + fund	3	50,44	13,43	-17,02	16,13	-7,72	-2,7	4,68	-11,28	0,64	46,6
		4	45,94	13,63	-18,84	14,2	1,1	-5,68	1,96	-5,62	-10,68	36,01
	ExEA. (10%)	5	33,74	6,94	-13,9	12	-4,27	-8,65	7,73	1,92	0,24	35,75
		6	38,18	8,83	-18,5	24,34	-0,32	-16,62	0,83	-8,39	3,82	32,17
	CoEA(10%)	7	21,77	21,31	-10,84	7,17	-4,89	-10,48	5,3	4,72	0,27	34,33
		8	13,56	24,85	-9,7	1,22	-6,27	-11,41	11,58	3,16	-0,17	26,82
	CoEA(30%)	9	45,9	21,77	-17,24	1,88	-2,17	-12,61	10,4	-7,2	-2,29	38,44
		10	11,15	33,1	-9,07	-1,75	0,9	-10,68	13,34	2,82	-5,11	34,7
	CoEA (10%) + fund	11	56,95	18,65	-16,69	17,97	-1,84	-4,94	0,44	-9,15	-14,63	46,76
		12	37,91	20,75	-15,73	19,75	-9,46	-16,45	6,71	-1,6	-10,28	31,6
H0	EA	13	44,15	12,87	-15,51	28,51	-2,55	-8,84	4,67	-1,57	-0,93	60,8
		14	49,17	10,22	-12,85	8,47	0,96	-3,38	4,25	1,58	-1,74	56,68
	EA + fund	15	36,3	7,65	-5,96	9,6	1,72	-3,08	2,26	1,96	-0,07	50,38
		16	49,12	9,03	-14,68	12,29	-3,46	-9,15	3,61	1,58	-4,59	43,75
	ExEA (10%)	17	49,33	11,09	-7,5	12,36	-2,02	-2,32	0,05	0,57	-0,3	61,26
		18	28,04	7,65	-2,94	6,01	1,82	-1,43	2,92	2,8	-0,07	44,8
	CoEA(10%)	19	44,15	12,94	-16,92	28,51	-3,55	-8,94	4,65	-0,83	-0,98	59,03
		20	49,39	13,46	-13,85	10,99	-1,49	-9,37	2,08	-0,05	0,95	52,11
	CoEA(30%)	21	47,7	14	-7,39	31,61	-0,64	-0,72	0,37	-0,51	0	84,42
		22	44,13	11,14	-7,87	11,51	-0,88	-1,7	0,8	-0,71	-0,15	56,27
	CoEA(10%) + fund	23	49,94	10,45	-16,35	21,91	-2,92	-8,76	6,03	1,38	-0,67	61,01
		24	39,32	16,93	-8,6	14,87	-0,07	-2,98	0,3	-0,53	-0,45	58,79
D0	EA	25	21,54	11,67	-9,12	13,26	-2,19	-14,55	5,22	1,91	1,39	29,13
		26	8,9	13,07	-3,79	6,07	5,42	-5,62	2,94	1,04	-1,31	26,72
	EA + fund	27	6,7	9,05	-8,73	12,25	3,41	-6,54	8,12	1,95	0,69	26,9
		28	22,35	2,83	-7,72	0,14	1,14	-10,09	13,57	4,33	-0,47	26,08
	ExEA (10%)	29	6,71	14,38	-8,08	-4,24	-1,97	-9,26	14,87	1,93	10,78	25,12
		30	9,04	5,47	-11,1	6,1	2,18	-10,27	4,11	3,6	9,92	19,05
	CoEA(10%)	31	29,42	8,8	-17,89	2,47	-2,38	-18,79	8,02	5,38	4,36	19,39
		32	7,5	4,17	-7,72	0,14	6,19	-10,07	11,53	-0,49	2,59	13,84
	CoEA(30%)	33	11,32	11,6	-6,25	-2,91	0,86	-10,33	10,17	0,52	0,99	15,97
		34	10,56	4,13	-9,77	0,14	-1,29	-10,81	14,03	3,86	4,17	15,02
	CoEA (10%) + fund	35	8,18	11,43	-8,69	-3,81	22,19	-5,82	8,61	0	5,02	37,11
		36	6,05	9,17	-8,73	7,2	1,65	-7,78	11,25	1,85	1,23	21,89
											Average:	39,13
											WIG:	**16,08**

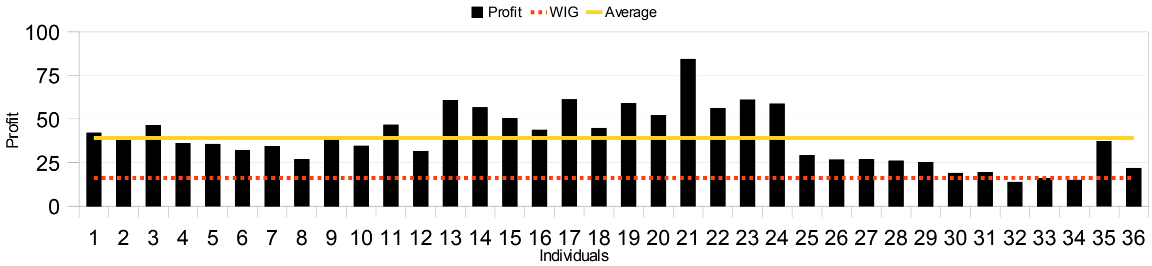

Fig 2: Test results (% return of initial capital) of best individuals gained in short time test

of +16,08% of the initial capital. Only three of the examined individuals gained lower profit ratio (Fig 2) and the averaged profit ratio is significantly higher (+39,13%).

The highest profits were turned by individuals trained on period *H0*. Their average profit (+57,44%) exceeds „buy-and-hold" strategy almost four times. Much lower, but also acceptable results achieved by individuals trained on uptrend period *U0* – their average gained profit was +36,93%. The worst results achieved by the individuals trained on downtrend *D0* – they earned *only* +23,02%. Let's note that two of the three highest results achieved individuals generated by *CoEA*.

B. Long term tests

To examine generated rules' profit in long term we used long term tests. The best individuals were selected from precious tests and tested in 7 periods as follows:

```
F1:  183 quoat., 179 shares, WIG  +24,09%
F2:  100 quoat,  245 shares, WIG   +1,04%
F3:   83 quoat.,  74 shares, WIG  +10,12%
F4:  150 quoat., 133 shares, WIG  +22,87%
F5:  154 quoat., 186 shares, WIG  +23,25%
F6:  131 quoat.,  80 shares, WIG  -25,33%
F7: 1835 quoat.,  32 shares, WIG +107,42%
```

The rules results gained on long term tests are presented in Table V . The profit ratio of „buy-and-hold" strategy gained in selected periods equals to +163,46%. In given test only 9 individuals reached the level of „buy-and-hold" profit strategy. The rest of strategies turned lower profits (it is average +113,64%).

The highest profits were turned by individuals trained in downtrend *D0*, where the average profit ratio equals to +179,98% and was about 25% higher than results of benchmark „buy-and-hold" strategy. Much lower profits were gained by individuals trained in uptrend period *U0*. Their average profit was +132,47%, but major contribution to this result was made by the individual no. 8 generated by *CoEA* in uptrend period F7 (it gained +604,51%). The lowest profits were turned by individuals trained in the horizontal trend period *H0*, where profit equals to +28,48% on average.

Let's analyze results details of the best individuals. Two of the three highest results were gained by individuals generated by the *CoEA* . The total profit ratio of the best individual (see no. 8 in Table V) equals to +624.92%. Unfortunately this individual managed well only in uptrend periods: it earned in *F3* (twice more than *WIG*), *F4* (slightly lower than *WIG*), *F5* (on an equal footing with *WIG*) and *F7* . In the uptrend period *F7* (as it is the longest of all examined periods) the given individual gained profit almost six times (!) higher than the growth of the whole market, and such profit has a tremendous effect on total individual profit value. All profits from the other uptrend periods were neutralized by losses from other trend periods. Even so we can say that the individual done quite well. Its trading rules:

BUY: *(ROC(5) < 0.97 AND (DMA(5,20) > 4.45 OR (RSI(5) < 9.86 OR RSI(5) < 71.12))) OR DMA(5,20) > 3.33*

SELL: *MACDO(12,26,9) > 5.81*

The individual buys shares while occurs fall in prices at least 3% (*ROC* lower than 1,0), but shorter moving average value is greater that long one value (positive *DMA* value) – it means that trend is going to change and it is strong signal to buy. If *DMA* does not satisfy condition there is analyzed *RSI* if it has a low value, what is connected to price decreasing in last sessions (what also generates a strong buy signal). A sell rule checks a moving average oscillator *MACDO* value and is interpreted as share price in last few sessions (a positive signal do sell).

The second individual (see no. 29 in Table V) that gained the highest profit (+426,27%) was generated by the *ExEA* . It can be noticed that this individual also earned money only in uptrend periods, where profits higher than *WIG* it turned in periods *F3* , *F5* and *F7* . In horizontal trend and downtrend periods it yields losses. Trading rules of this individual are given below:

BUY: ***INX < 38.50*** *AND ((RSI(10) < 2.40 OR RSI(5) <40.69) OR RSI(5) < 49.77)*

SELL: *(ROC(5) > 1.31 OR (RSI(10) > 65.16 OR MACDO(12,26,9) > 2.10)) OR ROC(10) < 0.40*

This individual also buys when price is relative low (*RSI* in 5 or 10 session is lower that 40,69 and 2,40). Unfortunately, the individual takes into consideration price of share, what can be connected to the unprofitable overfitting data effect. The sell rule checks *ROC* if prices are rising rapidly - about +31% within 5 session and if *RSI* is close to its "repurchase level" (it is possible a growth limit) and also

TABLE V. LONG TIME TEST RESULTS OF BEST INDIVIDUALS

Training Period	Algorithm type	No.	F1 (24,09%)	F2 (1,04%)	F3 (10,12%)	F4 (22,87%)	F5 (23,25%)	F6 (-25,33%)	F7 (107,42%)	Total Profit [%]
U0	EA	1	24,94	-23,41	10,07	4,72	30,28	-31,96	129,32	143,96
		2	-5,78	-18,36	8,06	16,28	22,76	-20,69	27,35	29,62
	EA + fund	3	60,54	-1,82	1,76	2,41	26,24	-21,54	124,23	191,82
		4	17,13	-6,11	-5,24	9,31	18,91	-34,05	-22,82	-22,87
	ExEA. (10%)	5	62,03	-7,53	6,9	0,83	46,79	-31,34	20,78	98,46
		6	-3,72	-13,65	17,53	-9,91	25,69	-33,15	0,66	-16,55
	CoEA(10%)	7	15,64	-12,7	2,36	3,55	30,75	-13,96	27,43	53,07
		8	**-0,36**	**-18,57**	**20,95**	**16,43**	**24,32**	**-22,36**	**604,51**	**624,92**
	CoEA(30%)	9	15,66	-24,7	-5,9	1,96	49,98	-24,23	122,77	135,54
		10	6,15	-16,32	5,66	18,87	39,81	-19,6	45,74	80,31
	CoEA (10%) + fund	11	61,31	3,12	-12,15	1,78	10,87	-39,01	61,67	87,59
		12	26,52	-6,05	-7,27	-1,92	112,25	-39,9	100,16	183,79
H0	EA	13	23,98	-1,41	6,09	10,63	38,54	-12,34	-24,97	40,52
		14	-2,15	-3,83	5,72	20,01	52,82	-30,74	2,18	44,01
	EA + fund	15	-4,85	-7,45	7,93	8,02	20,24	-20,93	9,81	12,77
		16	0,02	-4,6	9,33	4,09	19,37	-26,25	16,28	18,24
	ExEA (10%)	17	-3,98	-3,55	4,28	9,9	33,05	-19,73	-2,99	16,98
		18	2,32	-1,41	5,15	11,11	47,4	-8,56	-39,15	16,86
	CoEA(10%)	19	-12,49	-7,45	8,24	6,67	15,52	-21,08	79,16	68,57
		20	19,51	5,43	7,43	9,34	3,28	-24,25	1,96	22,7
	CoEA(30%)	21	18,77	4,76	-0,23	0,67	52,96	-12,11	1,06	65,88
		22	0,41	-1,71	3,33	10,04	35,98	-12,83	-29,37	5,85
	CoEA(10%) + fund	23	3,86	-7,45	8,36	16,86	21,62	-22,46	4,16	24,95
		24	-14,94	7,64	4,96	11,86	32,36	-13,22	-24,19	4,47
D0	EA	25	94,17	-15,5	14,24	0,7	55,48	-38,32	-39,83	70,94
		26	26,28	8,38	23,18	4,83	23,62	-23,96	136,94	199,27
	EA + fund	27	11,13	6,57	31,42	21,47	37,2	-27,47	128,65	208,97
		28	29,95	-15,36	15,01	11,54	33,58	-32,91	20,91	62,72
	ExEA (10%)	**29**	**19,73**	**-24,46**	**21,45**	**14,51**	**40,51**	**-26,02**	**380,55**	**426,27**
		30	27,79	-8,72	31,96	9,69	43,33	-31,4	106,99	179,64
	CoEA(10%)	**31**	**88,36**	**-21,2**	**22,47**	**0,51**	**58,72**	**-36,43**	**136,69**	**249,12**
		32	-4,6	-15,36	22,87	20,78	32,67	-30,75	187,28	212,89
	CoEA(30%)	33	19,46	-20,64	20,17	0,51	0,4	-29,98	167,87	157,79
		34	31,17	-16,17	19,66	11,42	31,1	-29,05	20,91	69,04
	CoEA (10%) + fund	35	-27,69	-20,88	14,54	-11,3	135,24	-2,98	20,22	107,15
		36	-3,79	-4,94	29,51	28,36	32,92	-24,7	158,59	215,95
									Average:	113,64
									WIG:	**163,46**

positive value of *MACDO* oscillator (the shorter moving average is up to the long one) informing about the price uptrend. The last condition (*ROC*(10)>0,40 - checks if price in last 10 session has decreased more than 60%) is discussable because can generate sell signal not in due time (our experiments have confirmed that).

The trade strategy of above individual it is strongly depended on share index value, what gives excessive adjustment to selected shares, what decrease its generalization ability and investing in other periods.

The third selected individual from the highest profits was turned by an individual (see no. 31 in Table V , generated by *CoEA*) earned +249,12%. This individual managed very well in uptrend periods (*F1* , *F3* and *F5*), but its results in periods *F2* and *F6* are considerable losses. In period *F4* in spite of market growth it wasn't able to yield any profit, but in the longest test period *F7* it turned a profit of +136,69% (in which market rose by +107,42%). Trading rules of this individual are as follow:

BUY: ***INX* < 3.48** *AND MACDO(12,26,9) < 10.21*

SELL: *RSI(10) > 66.27 OR ROC(10) > 1.37*

The above individual buys shares cheaper than 3,48 but only when *MACDO*>10,21 what makes this indicator usage rather discussable. The 'sell' signal is generated when share is is very close to its "repurchase level" (*RSI* says that it is possible a growth limit) or the price has increased more than +37% in the last 10 sessions. This individual also uses price value in its rule what may cause decrease a rule generalization quality (buys only 'cheaper' shares). On the other hand, this can be sort of trade specialization.

IV. SUMMARY

In the current stage of the project its results show that evolutionary algorithm it is an efficiency tool for useful knowledge discovery in stock market data. Our research results presented in previous section proves that it is possible to generate a trade strategy 'better' that benchmark "buy-and-hold" strategy.

The same tests results show also some of our approach disadvantages - standard deviation of profits is high, and we suppose that *EA* stacks in local optima and cannot search the problem solution space in efficiency way. The tremendous impact into *EA* efficiency has training period type selection, and we cannot say definitely which trade strategy is 'the best', so this is the main reason why strategies generated in given stage of project are rather risky to usage in real world applications by a beginner investor. It can be considered as tool for experienced investor as a inspiring and useful decision support tool. It is worth to noticing that the generated trade rules use indicators properly and in efficient way.

To increase efficiency of presented approaches we are going to include more specific technical indicators (such as oscillators or rankings, e.g. *RSI* to make it independent from current value) to discover more generalized trading rules. Also a set of indicators can be extended by some volume indicators (e.g. *OBV*) and removing a price indicators to avoid 'overlearning' and overfitting' the data (a trade rule is not general and buys only a 'favorite' company shares). Also promising way for improvement of our approach efficiency can be additional set of indicators that describes only trends in price/volume. Another proposition of indicator set extensions is including more fundamental indicators because presented version of our approach includes only five of them.

There are some inspiring *EA* research directions, such as specialized genetic operators including a local search or hill climbing method (in Baldwin effect), a rule pruning operator or hybrid approaches that links *EA* with other artificial intelligence tools such artificial neuronal networks or extend rule by fuzzy logic sets.

We suppose also that some our implementation assumptions causes indirectly *EA* search constraints, e.g. only 10 positions in portfolio and asking about buy/sell transaction of given share in alphabetical order (this promotes unnecessarily some shares). Thus limits company share selection and decrease efficiency portfolio management. Also in our approach we are not consider any risk aspects (such as company size, economy branch or other connections between companies) or portfolio diversification. They are very interesting further research directions.

References

[1] J. Arabas, „Wykłady z algorytmów ewolucyjnych" (eng. "Lectures on Evolutionary Algorithms"), WNT, Warszawa 2004 (*in polish*).

[2] G. Armano, M. Marchesi, A. Murru, „A hybrid genetic-neural architecture for stock indexes forecasting",Information Sciences: an Inter. Jour. vol.170 Issue 1, special issue: Computational intelligence in economics and finance, pp:3 – 33, 2005.

[3] P.-C. Chang, C.-H. Liu, J.-L. Lin, C.-Y. Fan, S. P. Celeste, „A neural network with a case based dynamic window for stock trading prediction", Expert Systems with Applications, 2009.

[4] P. Cichosz, "Systemy uczące się" (eng. "Self learning systems") WNT, Warszawa 2000 (*in polish*).

[5] D. B. Fogel, „Blondie 24: playing at the edge of AI", Morgan Kaufmann, San Francisco 2000.

[6] D. Goldberg, „Genetic algorithms in search, optimization and machine learning", Addision-Wesley, Boston, 1989.

[7] H.-J. Kim, K.-S. Shin, „A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets", Applied Soft Computing, 2007.

[8] J. Koza, „Genetic Programming", Cambridge, MIT Press 1992.

[9] R. J. Kuo, C. H. Chen, Y. C. Hwang, „An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network", Fuzzy Sets and Systems, Vol.118, Issue 1, 2001.

[10] X. Li, Z. Deng, J. Luo, „Trading strategy design in financial investment through a turning points prediction scheme", Expert Systems with Applications, 2009.

[11] Z. Michalewicz, „Genetic algorithms + data structures = evolution programs", Springer-Verlag Berlin Heidelberg, 1992.

[12] Z. Michalewicz, D. B. Fogel, „How to solve it: modern heuristics", Springer-Verlag Berlin Heidelberg, 2000.

[13] J. Morrison, „Co-evolution and genetic algorithms", Carleton University, Ottawa 1998.

[14] J. J. Murphy, „Technical analysis of the financial markets", Prentice Hall, 1999.

[15] P. B. Myszkowski, „Metody data mining w analizie giełdy", M.Sc. thesis Wrocław University of Techology, 2002 (*in polish*).

[16] J. Nenortaite, R. Simutis, „Development and Evaluation of Decision-Making Model for Stock Markets", Journal of Global Optimization, Vol.36 Issue 1, pp: 1-19, Springer , 2006.

[17] J. C. Ritchie, „Fundamental analysis", Irwin Professional Pub, 1996.