

Futures Trading: Design of a Strategy

Jan Zeman

Abstract—The paper describes the futures trading and aims to design the speculators trading strategy. The problem is formulated as the decision making task and such as is solved. The solution of the task leads to complex mathematical problems and the approximations of the decision making is demanded. Two kind of approximation are used in the paper: Monte Carlo for the multi-step prediction and iteration spread in time for the optimization. The solution is applied to the real-market data and the results of the off-line experiments are presented.

Keywords—futures trading, decision making

I. INTRODUCTION

The paper concerns with strategy design for futures contracts trading. A *futures contract* is a contract to buy (deliver) a specific amount of a commodity at a specific price and at specified date in the future. A *commodity* is any stuff for which there is a demand (e.g. basic resources and agricultural products such as iron ore, coal, sugar, wheat). Each futures contract has its own price, which reflects its actual demand on the market.

Speculators try to earn money by strategical buying and selling of contracts. They seem to play a game, where the participant bets on the increase or decrease of price in future. To design the trading strategy, speculators use various methods. The main streams are the fundamental analysis and the technical one. The fundamental analysis assumes that actual price does not reflect the real price, therefore bases predictions on analysis of the market state, actual news and activities of institutions. In contrast, the technical analysis deals primary by price curves to predict the further price behavior.

Classical investing methods based on fundamental analysis (e.g. value investing [1] or indexing [2]) serve primary for stock trading and the for long-time investment in terms of decades. The methods of technical analysis [3], unlike the fundamental one, provides profit in short-time, as it recommends actions more often, i.e. one action per week or month. However, there is no method of technical analysis, which results in profitable strategy working for decades. The viability of these approaches is about a year. Then, it should be completely revised.

Beside, the successful methods, if any, are not advertised everywhere and are kept in strict confidence. So up to the author's best knowledge, there is no known methodology how to design optimal strategy for speculators.

The paper employs Bayesian decision making to design speculator's trading strategy. Bayesian decision making is a well-developed methodology applicable to complex problems

with significant uncertainty and partial knowledge. Its functionality has been verified in natural, engineering and medical applications. It was proved [4], that designed strategy is non-dominated. The paper proposes the reformulation of trading task as decision making problem and subsequent solution. The complexity of explicit solution calls for suitable approximation that can be realized by existent tools.

The paper's outline is as follows. Section II introduces terminology of futures exchange, recalls main terms of decision making (DM) theory and reformulates futures trading problem as dynamic DM task. Section III contains approximation of DM. Section IV presents the experimental results obtained on real data. Section V addresses open questions as well as possible directions of the approach's improvement.

II. PRELIMINARIES

The following notations are used throughout: x^* denotes a set of all possible values of variable x . $|x|$ is absolute value of x . Conditioned expected value is

$$\mathcal{E}(x|y) = \int_{x \in x^*} x f(x|y) dx, \quad (1)$$

where $f(x|y)$ denotes conditioned probability density function (PDF). Index $t \in t^* = \{0, 1, 2, \dots, T\}$ labels discrete time instances. x_t denotes value of x at time t and $X_t^T = \{x_t, x_{t+1}, \dots, x_{T-1}, x_T\}$.

A. Trading futures

A futures contract gives the holder the obligation to buy or sell and the term *position* means a commitment to buy or sell a given amount of commodities. The basic types of position are distinguished: *short*, *long* and *flat*.

A *long position* yields a trader's benefit when the price increases, and trader's loss otherwise. This position refers to the situation when

- a trader buys an option contract that he has not already written (i.e. sold), he is said to be 'opening a long position'.
- a trader sells an option contract that he already owns, he is said to be 'closing a long position'.

A *short position* yields a trader's profit when the price decreases, and trader's loss otherwise. This position refers to the situation when

- a trader sells an option contract that he does not already own, he is said to be 'opening a short position'.
- a trader buys an option contract that he has written (i.e. sold), he is said to be 'closing a short position'.

A *flat position* denotes the state when no other type of position is active. Flat position means neither trader's profit nor trader's lose with any price change.

Jan Zeman is with the Department of Adaptive Systems, Institute of Information Theory and Automation, Prague, Czech Republic, e-mail: zeman@utia.cas.cz.

The opened position characterizes only potential gain, because the speculator holds the obligation, which price is changing. The potential profit becomes real by closing the position. The process from 'opening' up to 'closing' the position is called *trade*. The term trade is interesting to analyze designed strategy (Section IV-B).

The aim of trader is design such a strategy of positions selecting, which ensures trader's profit with minimal risk. The strategy design is based on prediction of price behavior and the quality of the strategy is very sensitive to small changes in strategy.

B. Decision making under uncertainty

Decision maker is either a human being or a device aiming to influence a part of the World he is interested in (so called *System*). The influence desired is expressed by DM aim. To reach this DM aim a decision maker designs and applies a sequence of *decisions* x_t . The decision typically influences the system, therefore decision maker works with respect to closed loop 'decision maker - system'.

All knowledge about system available to decision maker to design the decision x_t is called *experience* \mathcal{P}_t . *Ignorance* \mathcal{F}_t is knowledge about system unavailable to decision maker. *System behavior* consists of experience, decision and ignorance $\mathcal{Q} = (\mathcal{P}_t, x_t, \mathcal{F}_t)$.

The *strategy* is sequence of decision rules R_t , which maps the knowledge to decision.

$$R_t : \mathcal{P}_t^* \rightarrow x_t^*$$

The available knowledge grows with time, because it is extended each time step by new *system output* y_t and also by new decision x_t . The couple (x_t, y_t) is called *innovation* d_t . Thus, the experience could be expressed via $\mathcal{P}_t = (X_1^{t-1}, Y_1^t) = (D_1^{t-1}, y_t)$.

Gain express the success of reaching the decision maker aims with given decision making strategy. Gain is mapping of system behavior to real non-negative number $G : \mathcal{Q}^* \rightarrow [0, \infty]$, which is not causal and it is necessary to measure the potential strategy success. The *expected value* is functional which returns the value of the gain independent on ignorance for the given strategy and conditioned by experience. See [5] for explanation of the method.

The decision maker chose the decision $x_t \in x^*$ to maximize of expected value of the gain G conditioned by experience:

$$x_t = \arg \max_{x_t \in x^*} \mathcal{E}[G(\mathcal{Q}) | \mathcal{P}_t, x_t], \quad (2)$$

which is the idea based on principle of optimality - see [4].

C. Futures trades as DM task

This subsection reformulates futures trading task as a decision making problem, such as follows: The speculator as *decision maker* tries to earn at the exchange, which is the *system* of interest. The exchange provides a price of given contract as *system output* y_t . Speculator choses the short/long/flat position as his *decision* to maximize his profit (*gain*).

The system is exchange with one kind of futures contract. The system output y_t is a price of the contract. We design the

strategy for discrete time starting from 1, finishing by horizon T . The strategy starts and finishes with the flat position. We assume that our role on the market is so small, that our decision x_t does not influence the future price sequence Y_{t+1}^T .

The decision maker designs in each time t a number $x_t \in \mathcal{Z}$ as decision. The decision x_t characterizes traders position, i.e. $|x_t|$ characterizes count of contracts and $\text{sign}(x_t)$ characterizes the type of position 1 long, -1 short and 0 flat. The flat position at the beginning and at the horizon is expresses as: $x_0 = x_T = 0$.

The profit in time $t \in \{1, 2, \dots, T\}$ is expressed via:

$$g_t(D_{t-1}^t) = (y_t - y_{t-1})x_{t-1} - C|x_{t-1} - x_t|, \quad (3)$$

where $(y_t - y_{t-1})x_{t-1}$ is profit/loss caused by the change of price i.e. the decision x_{t-1} is profitable, when $\text{sign}(x_{t-1}) = \text{sign}(y_t - y_{t-1})$; C is normalized transaction cost and $|x_{t-1} - x_t|$ is change of position.

The gain for the whole trading horizon can be expressed as a sum of partial gains (3) over time $t \in t^*$. The gain function $G_t(\cdot)$ expresses the profit in time interval t, \dots, T :

$$G_t(D_{t-1}^T) = \sum_{k=t}^T g_k(D_{k-1}^k). \quad (4)$$

Easy to see, that the function $G_t(\cdot)$ is additive and backward recursive

$$G_t(D_{t-1}^T) = \sum_{k=t}^{t+h-1} g_k(D_{k-1}^k) + G_{t+h}(D_{t+h-1}^T), \quad (5)$$

which is valid for $h \in \{1, 2, \dots, T-t\}$.

To maximize the profit, the gain $G_1(\cdot)$ over the decisions x_1, \dots, x_T should be maximized. Using the optimality principle (see [4] for details), the optimal gain for $t \in \{1, \dots, T\}$ should be reached by maximization of admissible Bellman's function

$$\mathcal{V}_t(D_0^t, y_t) \equiv \max_{x_t, \dots, x_T} \mathcal{E}\left(G_t(D_{t-1}^T) \middle| D_0^t, y_t\right). \quad (6)$$

The expected value must be used to respect the fact that $G_t(\cdot)$ is function of the future data D_{t+1}^T . The expression (6) can be written as follows:

$$\begin{aligned} \mathcal{V}_t(D_0^t, y_t) = & \max_{x_t, \dots, x_{t+h-1}} \mathcal{E}\left[\sum_{k=t}^{t+h-1} g_k(D_{k-1}^k) \right. \\ & \left. + \underbrace{\max_{x_{t+h}, \dots, x_T} \mathcal{E}\left(G_{t+h}(D_{t+h-1}^T) \middle| D_0^{t+h}, y_{t+h}\right)}_{\mathcal{V}_{t+h}(D_0^{t+h}, y_{t+h})} \middle| D_0^t, y_t\right], \end{aligned}$$

where the second term is formally same as the original expression (6) with the time shift $t \rightarrow (t+h)$.

Then the Bellman's function holds the shape:

$$\begin{aligned} \mathcal{V}_t(D_0^t, y_t) = & \max_{x_t, \dots, x_{t+h-1}} \mathcal{E}\left[\sum_{k=t}^{t+h-1} g_k(D_{k-1}^k) \right. \\ & \left. + \mathcal{V}_{t+h}(D_0^{t+h}, y_{t+h}) \middle| D_0^t, y_t\right]. \quad (7) \end{aligned}$$

III. APPROXIMATION OF DECISION MAKING

Using the linearity of the expected value, the equation (7) results in:

$$\mathcal{V}_t(D_0^t, y_t) = \max_{x_t, \dots, x_{t+h-1}} \left[g_t(D_{t-1}^t) + \underbrace{\mathcal{E} \left(\sum_{k=t+1}^{t+h-1} g_k(D_{k-1}^k) \middle| D_0^t, y_t \right)}_{(*)} + \underbrace{\mathcal{E} \left(\mathcal{V}_{t+h}(D_0^{t+h}, y_{t+h}) \middle| D_0^t, y_t \right)}_{(**)} \right] \quad (8)$$

The term (*) models the expected multi-step gain to time $t+h-1$. Let assume that $x_t = x_{t+1} = \dots = x_{t+h-1}$. This assumption corresponds fact that often-trading systems lost almost of capital at transaction cost. On the other hand, the maximal generality can be hold by $h=1$. Using the assumption, equation (3) and properties of expected value, the term (*) can be written as follows

$$(*) = \left(\underbrace{\mathcal{E}(y_{t+h-1} | D_0^t, y_t)}_{\hat{y}_{t+h-1}} - y_t \right) x_t \quad (9)$$

The last unknown term is expected value of y_{t+h-1} . To express it using (1), the PDF of y_{t+h-1} is required. The PDF can be written in the parameterized form:

$$f(y_{t+h-1} | D_0^t, y_t) = \int_{\theta \in \theta^*} f(y_{t+h-1} | \theta, D_0^t, y_t) f(\theta | D_0^t, y_t) d\theta, \quad (10)$$

where θ is a vector of model parameters, $f(\theta | D_0^t, y_t)$ is the density of model parameters and $f(y_{t+h-1} | \theta, D_0^t, y_t)$ is the model of the of price y_{t+h-1} .

The assumed model is selected as n th order autoregressive and has following shape:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_n y_{t-n} + b + \epsilon_t, \quad (11)$$

where $\theta = (a_1, \dots, a_n, b)$ are model parameters and ϵ_t is white noise with distribution $N(0, \sigma^2)$. The one-step prediction based on equation (11) can be written in following shape:

$$\hat{y}_{t+1} = \mathbf{P} \underbrace{\begin{bmatrix} a_1 & \dots & a_{n-1} & a_n & b \\ 1 & & & 0 & 0 \\ & \ddots & & \vdots & \vdots \\ & & 1 & 0 & 1 \end{bmatrix}}_{\mathbf{M}(\theta)} \underbrace{\begin{bmatrix} y_t \\ \vdots \\ y_{t-n+1} \\ 1 \end{bmatrix}}_{\Psi_{n,t}},$$

where $\mathbf{P} = [1 \ 0 \ \dots \ 0]$ projects the first element from the vector $\mathbf{M}(\theta)\Psi_t$ and $\Psi_{n,t}$ is so called regressor of n length. Then, the multi-step prediction is expressed by

$$\hat{y}_{t+h} = \mathbf{P}[\mathbf{M}(\theta)]^h \Psi_{n,t}. \quad (12)$$

The PDF of model parameters $f(\theta | D_0^t, y_t)$ is estimated Bayesian filtration [6] and the results are first and second moments of PDF, which fully characterize it. With knowledge of the distribution $f(\theta | D_0^t, y_t)$, random samples $\theta_1, \dots, \theta_s$ can

be generated with the demanded distribution and multi-step prediction can be calculated:

$$\hat{y}_{t+h} = \frac{1}{s} \sum_{i=1}^s \mathbf{P}[\mathbf{M}(\theta_i)]^h \Psi_{n,t}. \quad (13)$$

The used algorithm corresponds with classical Monte Carlo algorithm.

Let approximate the term (**) of the equation (8). The main problem of calculating the term is backward character of equation (7), where the future value of Bellman's function $\mathcal{V}_{t+h}(\cdot)$ is needed to calculation the $\mathcal{V}_t(\cdot)$. To solve this problem the generalized shape of Bellman's function is demanded. The generalized shape can be obtained two ways: either expressing the generalized shape as formal solution of (7) or approximation by suitable shape.

Formal solution of equation (7) must be valid for all sequences D_0^T . However, this task is very complex and it seems impossible to find the generalized shape directly.

The approximation of Bellman's function is more promising way. Following design assumes, that Bellman's function shape does not vary. The approximation shape must be suitable for further computing, therefore the following shape has been chosen:

$$\mathcal{V}_t(D_0^t, y_t) \approx V_t(x_{t-1}, Y_{t-m+1}^t) \equiv p(x_{t-1})\Psi_{m,t}, \quad (14)$$

where $p(\cdot)$ is vector function $p: x^* \rightarrow \mathcal{R}^{1,m}$ and $\Psi_{m,t}$ is regressor. p is assumed to be independent on time t . Then, the expected value (**) can be written as follows:

$$\mathcal{E} \left(\mathcal{V}_{t+h}(D_0^{t+h}, y_{t+h}) \middle| D_0^t, y_t \right) = p(x_{t-1}) \mathcal{E} \left(\Psi_{m,t} \middle| D_0^t, y_t \right), \quad (15)$$

where expected value of regressor $\mathcal{E}(\Psi_{m,t} | D_0^t, y_t)$ contains multi-step predictions designed above.

Let focus at design function $p(\cdot)$. The support of $p(\cdot)$ is x^* , which is the bounded subset of whole numbers $x^* \subset \mathcal{Z}$. Therefore, the function is characterized by matrix $k \times m$, where m is dimension of regressor $\Psi_{m,t}$ and k is count of values x^* . Moreover, the set x^* is optional parameter of trading system chosen by user.

Substitution of (14) into equation (7):

$$p(x_{t-1})\Psi_{m,t} + e_t = \max_{x_t, \dots, x_{t+h-1}} \mathcal{E} \left[\sum_{k=t}^{t+h-1} g_k(D_{k-1}^k) + p(x_{t+h-1})\Psi_{m,t+h} \middle| D_0^t, y_t \right], \quad (16)$$

where e_t is a non-preciseness generated by the approximation. The equation (16) holds for each t and h .

Let denote $X_{opt}^n = \{x_1^o, \dots, x_n^o\}$ the best strategy at known data sequence Y_0^n . It is experimentally proven that: *For each dataset Y_0^T , there is a finite number $q \in N$ that for each $n \in \{q+1, \dots, T\}$, sequences X_{opt}^n and X_{opt}^T have first $(n-q)$ elements the same.*

Using this claim in time t , the optimal strategy X_{opt}^t has first $(t-q)$ actions identical as the best strategy X_{opt}^T designed

at whole sequence including the ignorance. Therefore, for first $\tau \in \{1, \dots, t - q\}$ can be equation (16) written as follows:

$$p(x_{\tau-1}^o) \Psi_{m,\tau} + e_{\tau} = \sum_{k=\tau}^{\tau+h-1} g_k((D_{opt})_{k-1}^k) + p(x_{\tau+h-1}^o) \Psi_{m,\tau+h}, \quad (17)$$

where $D_{opt} = (Y_1^t, X_{opt}^t)$. The expected value dismissed, because all used variables are known, maximum was realized by using the elements of the best strategy at whole data sequence. The obtained equation system for $\tau \in \{1, \dots, t - q\}$ contains function $p(\cdot)$ and other terms are known or calculable. Thus, the system can be used for estimation the values characterizing the function $p(\cdot)$.

The system enlarges about one equation in each time step and due to approximation of Bellman's function, could lost the solubility, when count of equations overgrows the count of free parameters characterizing $p(\cdot)$. Thus, the least squares method was used, with minimization over the non-preciseness:

$$\min_{p(\cdot)} \sum_{\tau=1}^{t-q} e_{\tau}^2, \quad (18)$$

where e_{τ} is expressed from equation (17).

IV. EXPERIMENTAL PART

This section describes the experimental setup, data and results obtained. The trading strategy is designed at discrete time $t \in \{1, 2, \dots, T\}$. The time step $[t, t + 1]$ corresponds with interval of 24 hours.

The data used as Y_0^T are so-called *close prices*, which are collected once a day. It is the last price, when the exchange closes trading. The economic specialists grant that close price is the most stable price. The transaction costs C were defined to the same values as at the real exchanges.

The general design presented above does not specify the restriction to decision x_t . In real market, the restrictions depend on the trader's account, as traders must own money to buy or sell contract at the exchange and the range of contracts to position is limited by owned money. Following values of decision $x_t \in \{-1, 0, 1\}$ were used for our academic experiments.

The order of model (see (11)) is set to $n = 2$, because this value generated best values in the previous research. Predictions are generated by Monte Carlo method. The count of Monte Carlo samples s (13) is chosen dynamically: the decision is final, when it is not influenced by new Monte Carlo samples. Bellman's function (14) is approximated by regressor of length $m = 5$.

A. Used data

The available dataset contains 35 price sequences for the experiments. The sequences contain prices for more than 15 years, i.e. about 3900 trading days in each sequence. The experiment set is too wide to present all results here, therefore the following five futures contracts were chosen as reference markets: Cocoa - CSCE (CC), Petroleum-Crude Oil Light (CL), 5-Year U.S. Treasury Note (FV2), Japanese Yen

- FOREX (JY) and Wheat - CBT (W). The reference markets were chosen by economic specialist to include all typical kind of markets - i.e. cocoa and wheat are typical agriculture product, petroleum-crude oil is mined material, Japanese Yen is typical foreign currency and treasury note stands for bond markets.

B. Results

There are many ways, how evaluate the quality of designed strategy. The main criterion is net profit $P(1, T)$. The net profit over the time t_1, \dots, t_2 is calculated by:

$$P(t_1, t_2) = \sum_{k=t_1}^{t_2} g(D_{k-1}^k). \quad (19)$$

Secondary criteria are gross loss (sum of the profit over lost trades), gross profit (sum of the profit over won trades) and count of winning and losing trades.

The main non-quantitative pointer is the plot of cumulative gain depending on time:

$$G_c(t) = P(1, t) = \sum_{k=1}^t g(D_{k-1}^k), \quad (20)$$

which increases, in the ideal profitable case.

The results overview is in Table I. Results in upper part of table contains values in \$1000 USD and the lower part contains count of trades. (A trade starts in time t_1 with choosing non zero decision x_{t_1} and finishes by choosing another value x_{t_2} in time t_2 .)

The reached results are quite good. Three markets were profitable, two have net profit around zero.

CL, FV2 and W are profitable and their cumulative gain steady increases. The results of FV2 and W are depicted at Fig. 1 and Fig. 2. The interesting is behavior of the system W after time 1500, where the price upward trend finishes and then is changed by downward trend. The system in this period starts oscillating and then sets again the profitable position. But in the oscillation between profitable intervals, the cumulative gain decrease due to transaction cost.

Cocoa (CC) and Japanese Yen (JY) are not as profitable as other markets, but the cumulative gain levels off around the zero. As can be seen from gross profit and gross loss, the difference between profitable and non-profitable trades is minimal, therefore this two markets are said to be around zero.

There are not any non-profitable market in the presented set.

TABLE I
 RESULT OVERVIEW

	CC	CL	FV2	JY	W
Net profit	-4.28	36.78	26.17	5.07	26.33
Gross profit	82.54	292.86	75.25	339.74	100.27
Gross loss	-86.82	-256.08	-49.08	-334.68	-73.93
Transaction cost	-6.16	-12.94	-2.92	-8.62	-4.68
Trades	308	647	146	431	234
Wining trades	165	356	74	227	129
Losing trades	143	291	72	204	105

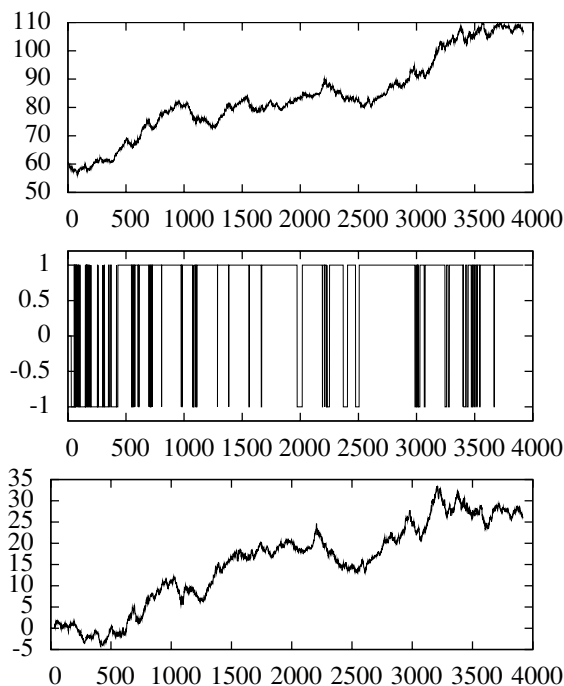


Fig. 1. Results on FV2: 1st chart is price y_t scaled to points (1 pt = \$1000 USD), 2nd chart depicts decisions x_t , 3rd chart depicts cumulative gain $G_c(t)$ with the same scale as is used in price (1 pt = \$1000 USD).

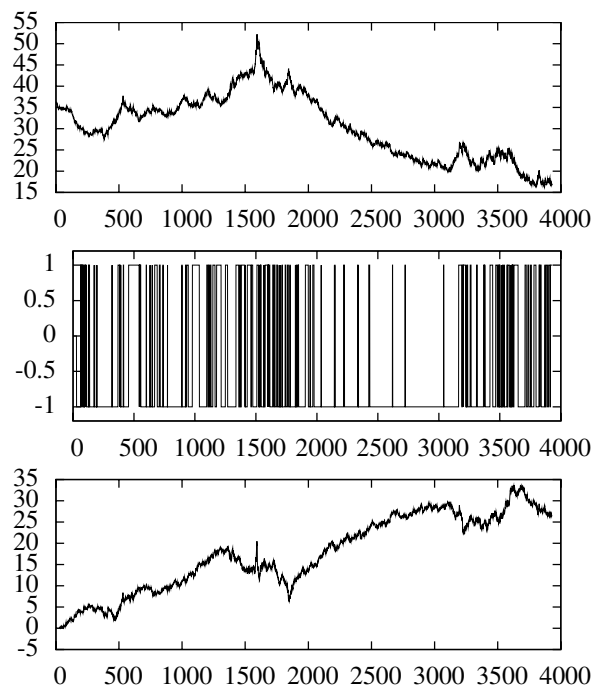


Fig. 2. Results on W: 1st chart is price y_t scaled to points (1 pt = \$1000 USD), 2nd chart depicts decisions x_t , 3rd chart depicts cumulative gain $G_c(t)$ with the same scale as is used in price (1 pt = \$1000 USD).

V. CONCLUSION AND FUTURE WORK

The design of the speculator strategy is presented. The strategy is designed using the Bayesian decision making. The reformulation of the problem as decision making task leads to generally unsolvable equations, therefore the approximations are used. To find multi-step predictions, the Monte Carlo method is used. And iteration spread in time, with approximation of the Bellman's function, is used for optimization. The strategy design is tested at the real-prices data and the basic result are presented.

This paper presents basic design and forms basis for further approaches. The standalone basic solution supports the research team, because researchers could deal only their special part of system to improve. The functional basis gives quick feedback, whether the improvement of special method leads to profit or not.

The further research should be on two directions: prediction model and optimization. The model should be extended about the additional information channels, because the close price is not enough to characterize all necessary market properties. Following information is available and will be included: additional price statistics, count of trades per time and commitment of traders information. Therefore the multidimensional model should be designed first as autoregressive model and then as general model if it would be necessary. The prediction, connected with model, should be enrich by working with PDF, not only with first moment of PDF, as is presented here.

The optimization consists of two closely related parts: approximation of Bellman's function and optimization. The Bellman's function used in present optimization demands for improvement, the least squares should be changed to

weighted least squares. The optimization should cover the multi-dimensional model with the multi-step prediction.

ACKNOWLEDGMENT

The author would like to thank grants GACR 102/08/0567, MŠMT 2C06001 .

REFERENCES

- [1] B. Graham, *Security Analysis*. New York: McGraw Hill Book Co., 1934, 4. ISBN 0-07-144820-9.
- [2] J. Bogle, *The First Index Mutual Fund: A History of Vanguard Index Trust and the Vanguard Index Strategy*. Bogle Financial Center.
- [3] "Stockcharts (<http://www.stockcharts.org/>)."
- [4] D. Bertsekas, *Dynamic Programming and Optimal Control*. Nashua, US: Athena Scientific, 2001, 2nd edition.
- [5] I. Nagy, L. Pavelková, E. Suzdaleva, J. Homolová, and M. Kárný, *Bayesian decision making*. Prague: CAS, 2005.
- [6] M. Kárný, B. J., T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma, and L. Tesař, *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. London: Springer, 2005.



Jan Zeman received a MSc. degree in Software Engineering at the Faculty of Nuclear Sciences and Physical Engineering (Czech Technical University) in Prague in 2007. He currently studies Mathematical Engineering at the same faculty and works on his PhD. thesis at the Department of Adaptive Systems of Institute of Automation and Information Theory of the Czech Academy of Sciences under supervision of Tatiana Valentine Guy.