Scientific Research of the Institute of Mathematics and Computer Science, 2(10) 2011, 75-86

SIMULATION APPROACH TO OPTIMAL STOPPING IN SOME BLACKJACK TYPE PROBLEMS

Andrzej Z. Grzybowski

Institute of Mathematics, Czestochowa University of Technology, Poland
azgrzybowski@gmail.com

Abstract. In the paper, an unbounded blackjack type optimal stopping problem is considered. A decision maker (DM) observes sequentially the values of an infinite sequence of nonnegative random variables. After each observation, the DM decides whether to stop or to continue. If the DM decides to stop at a given moment, the obtains a payoff dependent on the sum of already observed values. The greater the sum, the more the DM gains, unless the sum exceeds a given positive number. If so, the decision maker loses all or part of the payoff. It turns out that under some elementary assumptions the optimal stopping rule (OSR) for such a problem has a very simple, so-called threshold form. However, even in very simple cases, the value of the problem has no closed analytical form. Therefore, it is very hard to evaluate the value directly. Thus, in order to find the relationship between the problem design parameters and the value of the problem, is proposed studying the relation via Monte Carlo simulations combined with regression analysis The same approach is adopted to examine the OSR risk characteristics.

Introduction

The “blackjack type problem” (BTP) models a class of optimal stopping decision tasks in which the decision maker observes sequentially the values of a given, maybe infinite, sequence $X_1, X_2, \ldots, X_N, \ldots$ of nonnegative random variables. After each observation, the decision maker (DM) decides whether to stop or to continue. If the DM decides to stop at moment k, he/she obtains a payoff dependent on the sum $X_1+\ldots+X_k$. The greater the sum, the more the DM gains, unless the sum exceeds a given number T **-** a limit given in the problem. If so, the DM loses all or part of the payoff. Such problems can represent various real world situations, which can be observed in engineering, economics, finance or social life, see e.g. [1, 2]. To illustrate the class of problems, let us consider a problem of *loading a device with a limit of load bearing capacity*. Many types of machines (trucks, cranes etc.) or other engineering structures (such as dams, roofs, bridges, computer servers) may be subjected to excessive overloads resulting in possible breakage of the mechanism or structure. Assume a DM observes the loading process of such a device. During the loading process, the load is increased in random steps, as for example during a flood (a dam) or heavy snowfall (a roof). Assume that the limit of the load bearing capacity of the device is given. After each observation, the DM decides whether to stop or to continue the process of loading. The DM wants the device to bear as much load as possible. However, on the other hand, if the limit of the load bearing capacity is crossed, then the gain for the DM is dramatically decreased.

The name of the class of optimal stopping problems is taken from one of the most popular casino table card games. *Blackjack type games* are played on a points system that gives numeric values to every card in a single deck of playing cards. The cards are given to a player sequentially until he decides to stop. The score is the sum of the values in his hand. The player with the highest total score wins as long as it does not exceed a given limit. If a player's cards exceed the limit, then the player loses and his/her bet is taken by the dealer.

Optimal stopping problems form a class of optimization problems with a wide range of applications in mathematical statistics, engineering, industry, economics, and mathematical finance. The most interesting include e.g. job-search and house-hunting problems, see e.g. [3-6], engineering and computer systems maintenance and/or management [7, 8], the pricing of perpetual American options as well as the optimal timing to invest in a project or capitalizing an asset [9-12].

In the theory of optimal stopping, see e.g. [3, 4], the solution of any optimal stopping problem consists of the optimal stopping rule (OSR) and the value of the problem, i.e. the greatest expected payoff possible to achieve. In the case of a finite horizon, a solution for BTP satisfying some general assumptions is given in [1]. It appears that the OSR has a relatively simple structure. However, the dependence between the expected gain and the design parameters of the problem is rather complex. Even more complex is the relation between these parameters and the value of the problem in the case of an infinite horizon. Another important problem is to describe the relation between various risk characteristics of a given OSR and the parameters of the stopping problem. There is no analytical expression relating the design parameters of the decision problem to the corresponding performance characteristics of the decision rule. Usually in the case when the relationship between some dependent and the independent variables is extremely complex or unknown, the Monte Carlo simulations approach can be adopted, see e.g. [2, 13-15]. However, the Monte Carlo methods allow us to solve a specific given problem rather than to obtain some general expressions describing the relation in which we are interested. Thus in order to obtain some more general results we propose combining the Monte Carlo method with regression analysis which enables us to estimate and express analytically the relationship which we are going to study with the help of computer simulations.

The paper is organized as follows. In the next section we formally state a general BTP and recall some important definitions from the theory of optimal stopping. In Subsection 1.3, we define the considered risk characteristics and in Subsection 1.4, we describe a specific BTP which will be studied in detail. In Subsection 2.1, we describe the Monte Carlo experiment which we use in order to obtain data containing the information about the relations between the risk characteristics and design parameters of the BTP. Next, we adopt regression analysis in order to obtain approximations of the analytical expressions relating the value of the problem as well as the risk characteristics with the design parameters of the considered BTP .

1. Formal statement of problem

The formal model for the class of problems we consider in the paper is the following. Let X_1, X_2, … be an infinite sequence of random variables. A DM observes sequentially the values of the variables and decides whether to stop or to continue. If the process is stopped at moment k, the DM gains value $W(y+\sum_{i=1}^{k} X_i)$, where W is a given real function and $y \geq 0$ is the initial state of the process. Function W is positive and nondecreasing on the interval $(0, T]$ and is nonincreasing for arguments greater than T. Such problems will be called *blackjack type problems* (BTP) if the random variables are *nonnegative* and payoff function W achieves its *only* maximum for $y+\sum_{i=1}^{k} X_i = T$.

Our task is to find a stopping rule which maximizes the expected payoff for a decision maker.

1.1. Optimal stopping theory - necessary definitions and results

Before we present the problem considered in the paper we need to present some necessary formal definitions from the theory of optimal stopping. They can be found e.g. in [3, 4].

Let X_1, X_2, … be a sequence of independent random variables. Let $\mathcal{F}_n$ denote σ-algebra generated by random variables $X_1, X_2, \ldots, X_n$ in an underlying probability space (Ω $\mathcal{F}$,P) *A stopping rule* is a random variable τ with values in a set of natural numbers such that $\{\tau = n\} \in \mathcal{F}_n$ for n = 1,2,… and P($\tau < \infty$) = 1. Let **M**(n) be a class of all stopping rules τ such that P($\tau \leq n$) = 1.

Let $(Y_n, \mathcal{F}_n)$, n = 1,2,…, be a homogenous Markov chain with values in a state space ($\mathcal{Y}$, $\mathcal{B}$). Let $W: R_+ \to R$ be a Borel measurable function whose values $W(y)$ will be interpreted as the gain for a DM when chain $(Y_n, \mathcal{F}_n)$ is stopped at state y. Assume that for a given state y and for a given stopping rule τ, expectation $E_y W(Y_\tau)$ = = $E(W(Y_\tau)|Y_1 = y)$ exists. Value $E_y W(Y_\tau)$ is the mean gain corresponding to the chosen stopping rule τ.

Let us define a function V_n by the equation:

$$V_n(y) = \sup_{\tau \in M_W(n)} E_y W(Y_\tau) \tag{1}$$

where $\mathbf{M}_W(n)$ is the set of all stopping rules belonging to **M**(n) for which expectations $E_y W(Y_\tau)$ are larger than $-\infty$ for all $y \in \mathcal{Y}$. Value $V_n(y)$ is called a *value* of the problem of optimal stopping when the *initial state* of the process is y and the *boundary* (horizon) for the possible number of steps is N.

Stopping rule $\tau^* \in \mathbf{M}_W(n)$ which for all $y \in \mathrm{Y}$ satisfies the condition

$$E_y W(Y_{\tau^*}) = V_n(y) \tag{2}$$

is called an *optimal stopping rule* in class $\mathbf{M}_W(n)$.

Now let us consider an unbounded problem. Let $\mathbf{M}_W$ denote the set of stopping rules satisfying the conditions: $P(\tau < \infty) = 1$ and $E_y W(Y_\tau) > -\infty$ for all $y \in \mathrm{Y}$. The value of such a stopping problem is denoted by $V(y)$ and the stopping rule which satisfies a condition analogous to (2), with V_n replaced by V, is called an optimal one in class $\mathbf{M}_W$.

1.2. Certain unbounded blackjack type problems and their solutions

The following proposition providing us with a solution for bounded BTP is proved in [1].

Proposition 1. If there exists real number t^*, $0 < t^* < T$, such that

$$W(y) < \mathrm{E}_y W(Y_1) \text{ for } 0 \le y < t^*$$

and (3)

$$W(y) \ge \mathrm{E}_y W(Y_1) \text{ for } y \ge t^*$$

then OSR τ_n^* in class $\mathbf{M}_W(n)$ for the BTP is given by

$$\tau_n^* = \min\{1 \le k \le n : Y_k \ge t^*\} \tag{4}$$

Value $V_n(y)$ of the problem can be calculated for $y < t^*$, with the help of the following recursive equation:

$$V_n(y) = \int_0^{t^*-y} V_{n-1}(y+x) f(x)dx + \int_{t^*-y}^{\infty} W(y+x) f(x)dx \quad n = 2,\dots,N \tag{5}$$

with the initial condition $V_1(y) = \int_0^{\infty} W(y+x) f(x)dx$.

We see that the above OSR is of the so-called threshold type. Such OSRs are especially practically interesting because of their very simple structure, compare [16-19].

Now let us consider an unbounded version of such a problem. It follows from well -known theorems that under some assumptions, unbounded optimal stopping rule τ^* can be approximated by bounded optimal stopping rules τ_n^*. One of these theorems, see Th.11, p. 77 in [4], states that if the payoff function is bounded, then

$$\tau^* = \lim_{n\to\infty} \tau_n^*$$

and

$$V(y) = \lim_{n\to\infty} V_n(y) \tag{6}$$

It results from the definition of the BTP that without the loss of generality, we can assume that the payoff function is bounded. Thus the above mentioned theorem yields the following proposition.

Proposition 2. If conjunction (3) is satisfied, then the OSR in class $\mathbf{M}_W$ for the BTP is given by

$$\tau^* = \min\{1 \leq k : Y_k \geq t^*\} \tag{7}$$

Condition (3) is fulfilled in many practically interesting problems, for examples see [1]. One of such problems will be considered in detail in the sequel.

1.3. Important characteristics of OSR

In the situation where we deal with decision making under uncertainty, the most important for the DM features of any decision rule are the expected payoff and the risk characteristics.

It results from the two above propositions that value $V_n(y)$ of a bounded problem can be computed with the help of recursive equation (5). However, usually the calculations are extremely arduous, even if we make use of some symbolic manipulation software, such as *Maple*, *Mathematica* or *Maxima*, see [1, 2]. In the case of an unbounded problem, there is not even one recursive formula to calculate the value. The same problem is connected with the risk. The theory of optimal stopping hardly provides us with any results devoted to any risk measures connected with the optimal stopping rule.

In general decision theory there are two basic types of risk concepts:

– risk connected with the variability of results around a specific value of payoff
– risk connected with the possibility of occurrence of undesired results.

For the BTP, two risk characteristics reflecting both above risk concepts were proposed in [2]. Let Z be the random payoff connected with optimal stopping rule τ^*, i.e. $Z = W(Y_{\tau^*})$. Let σ_Z denote the standard deviation of optimal payoff Z. The following risk measures connected with rule OSR τ^* will be considered in the sequel:

– ratio SV of standard deviation of random payoff to expected payoff, i.e. $SV = \sigma_Z/V_N(0)$
– probability of failure PrF, i.e. probability that the process under control will cross limit T.

In the sequel we deal with the problem of modeling the expected payoff as well as the two risk characteristics for the unbounded BTP. We combine the Monte Carlo method with regression analysis to estimate and express analytically the relationship between the design parameters of the BTP and the indicated characteristics of the OSR. The BTP we will study in detail is described in the next subsection.

1.4. Blackjack type problem with linear payoff and exponential step

In the sequel, the following BTP will be considered in detail. Let the DM observe a sequence of i.i.d. random variables having an exponential distribution with the density function:

$$f(t) = \frac{1}{\lambda}\exp(-\frac{t}{\lambda})\mathbf{1}_{[0,\infty)}(t), \quad \lambda > 0 \tag{8}$$

Therefore, in this problem the DM approaches limit T with exponential steps of an average length λ.
Let payoff function W be given by the following equation:

$$W(y) = \begin{cases} B \cdot y, & y \le T \\ 0, & y > T \end{cases} \tag{9}$$

with $B > 0$.

According to formula (8), the DM obtains a positive payoff which is proportional to state y of the process, unless the state is greater than limit T. If so, then the player gains 0.

It was shown in [1] that such a problem satisfies the condition given in Proposition 1 with

$$t^* = T - \lambda \ln(1 + \frac{T}{\lambda}) \tag{10}$$

the OSR τ^* given by (7) tells us to continue the observation as long as the sum of the initial state and already observed values does not exceed the above given value t^*.
This particular subclass of BTP models an important practical decision task connected with the theory of mass-service and called *service with work time limits*, see [1, 2].

2. Monte Carlo simulations and regression models

In this section, we describe the Monte Carlo experiment and present the results of the regression analysis applied to the obtained data.

2.1. Monte Carlo experiment

The idea of the Monte Carlo simulation is to draw sample $\{Z_i\}_1^m$, i.e. a realization of stochastic process $\{Z_1, Z_2, ...Z_m\}$ composed of independent and identically distributed random variables having the same distribution as random optimal payoff $Z = W(Y_{\tau^*})$. Let f be any Borel function for which expected value $Ef(Z)$ exists.

By the strong law of large numbers, average $\bar{f}_m = \frac{1}{m}\sum_{i=1}^{m} f(Z_i)$ will almost surely (*a.s.*) converge to E*f(Z)*. In particular, when *m* tends to infinity, we have

$$\bar{Z}_m = \frac{1}{m}\sum_{i=1}^{m} Z_i \xrightarrow{a.s.} E(Z) = V_N(y)$$

$$S_m^2 = \frac{1}{m}\sum_{i=1}^{m} (Z_i - V_N(y))^2 \xrightarrow{a.s.} \sigma_Z^2 \tag{11}$$

and

$$\frac{M}{m} = \frac{1}{m}\sum_{i=1}^{m} \mathbf{1}_{(-\infty,a]}(Z_i) \xrightarrow{a.s.} \Pr(Z \le a)$$

In the latter expression, *M* is the number of the values in the Monte Carlo sample which are not greater than *a*.
In our Monte Carlo simulations, the realizations of random optimal payoff *Z* are generated directly according its definition with the help of the following procedure:

```
Set z=y;
While y≤t* Do Set z=z+REX(λ)
If y<T Set W=B·z
Else   Set W=0
Return W
```

In the above procedure, function REX(λ) returns a pseudorandom number generated according to the exponential probability distribution having the density function given by (8), *t** is given by (10). We use the procedure to estimate the values of the problems for various design parameters as well as other statistics characteristising the performance of the OSR.

2.2. Design parameters of the problem

The design parameters of a given BTP as stated in section 2.1 are the following: limit *T*, parameter λ determining the step probability distribution, and payoff function parameter *B*. Let us assume that initial state *y* of the process equals 0 and let us confine ourselves to these situations where the value of problem *V*(0) is positive. It reflects the case where the optimal stopping rule tells the decision maker to make at least one observation.

It appears, see [2], that it is very convenient to consider a parameter *K* which equals ratio T/λ and can be interpreted as the average number of steps needed to cover distance *T*. It allows us to obtain more general results. Because the optimal stopping rule is independent of *B* and the expected value of the payoff as well as the value of the problem are linear functions of *B*, we assume in the sequel that *B* = 1. What is more, to obtain an even more general description, we model the ratio

$$R = \frac{V(0)}{T} \tag{12}$$

instead of modeling the problem value alone. Thus finally, we have one independent variable for our models - parameter K.

2.3. Monte Carlo estimation of the value of the problem

It was shown in [2] that the Monte Carlo approximations of the value of the problem in the case of bounded versions of the BTP are very accurate. The average relative error of these approximations was about 0.3%, see [2]. Thus one may expect the same in the case of the unbounded BTP. Let V^{MC}, R^{MC} denote the Monte Carlo estimate of the values V and R respectively. We compute the estimates of V^{MC} and R^{MC} for the values of parameter K changing in interval [1, 30]. In our Monte Carlo simulations we assume m = 10 000, compare (11), and for each number K limit T is chosen at random from interval [50, 250]. Next we adopt the regression analysis to obtain an analytical model relating ratio R given by (12) and parameter K. The resulting model has the following form:

$$R(K) = \begin{cases} \beta_0 + \beta_1 / K + \beta_2 K + \beta_3 K^2, & \lambda \le g \\ \beta_4 + \beta_5 / K + \beta_6 K + \beta_7 K^2, & \lambda > g \end{cases}$$

with the following *least squares* (*LS*) estimates b_i for the unknown coefficients β_i, $i = 0,\dots,8$:

$b_0 = 31.1854$, $b_1 = -11.2167$, $b_2 = 11.5469$,
$b_3 = -0.917678$, $b_4 = 79.6558$, $b_5 = -94.3583$,
$b_6 = 0.6659$, $b_7 = -0.00867$, $g = 4.8924$.

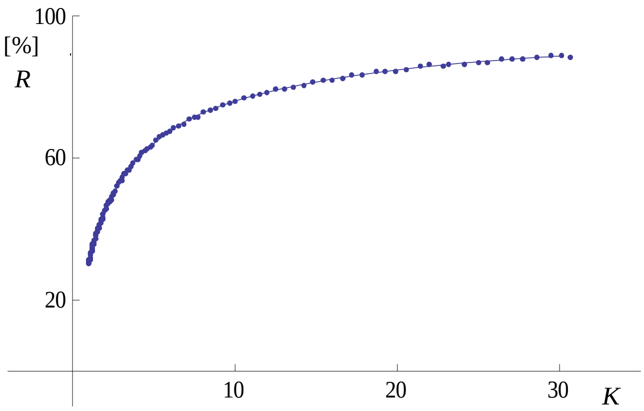

Fig. 1. Graph of model for ratio R (in %) as function of K and Monte Carlo approximation of values obtained for $K \in [1, 30]$ (dots). Regression values and their approximations are almost the same - they can hardly be distinguished in the figure

Figure 1 shows both the data obtained by the Monte Carlo experiment, as well as the graph of a model for ratio *R* estimated on the basis of the Monte Carlo data. We see that the model values and the Monte Carlo approximations can hardly be distinguished.

Now, to study the quality of the approximations we compute the average relative error *RE* between our model and of the Monte Carlo approximations.
The formula for *RE* is as follows:

$$RE = \frac{1}{N}\sum_{i=1}^{N}\left|R(K_i) - R^{MC}(K_i)\right| / R(K_i) \tag{13}$$

To compute RE, we generate another Monte Carlo sample (called in the sequel a *validation* set) containing $N = 4000$ records. The value of *RE* obtained for our data is 0.00288, and its value confirms that the regression model is really good.
In the next part of the paper, we adopt this approach to build models for the risk characteristics of the optimal stopping rule.

2.4. Regression models for risk characteristics

Model risk characteristics are developed on the simulations described in the previous subsection.

First we present the model for ratio *SV*. We assume the following form of the regression function:

$$SV(K) = \begin{cases} \beta_0 + \beta_1 / K + \beta_2 K + \beta_3 K^2, & \lambda \le g \\ \beta_4 + \beta_5 / \lambda + \beta_6 \lambda, & \lambda > g \end{cases}$$

Based on the Monte Carlo data, we obtain the following *LS* estimates b_i of unknown coefficients β_i, $i = 0,\ldots,6$:

$b_0 = 0.6948$, $b_1 = 0.508183$, $b_2 = -0.11354$, $b_3 = 0.00980$, $b_4 = 0.212579$,
$b_5 = 1.3666$, $b_6 = -0.00260$, $g = 4.9530$

The model function graph along with the data is presented in Figure 2.

To check the usefulness of the regression model, we compute the *RE* given by (13) (with *R* replaced with *SV*) based on the validation set. The average relative absolute error *RE* equals 0.63%. It confirms good quality of the regression model.

Figure 3 illustrates the dependence between the probability of failure *PrF* and parameter *K*. The continuous line represents the graph of the regression model obtained in our studies. It has the logistic form given by the formula:

$$\mathrm{PrF}(K) = \frac{\mathrm{Exp}(PF(K))}{1 + \mathrm{Exp}(PF(K))}$$

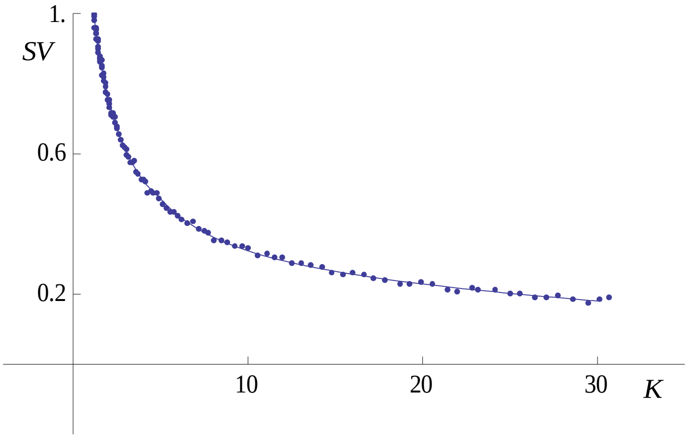

Fig. 2. Graph of estimated regression function *SV* (continuous line) and Monte Carlo estimates for *SV* (dots) when $K \in [1, 30]$, and $m = 10000$. Parameter *T* was chosen at random from interval [50, 250]

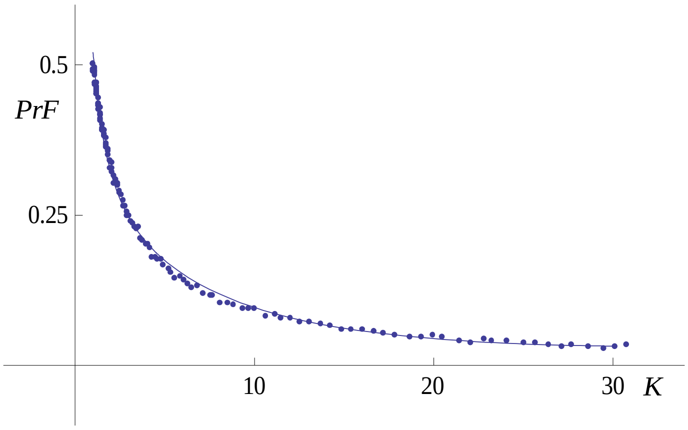

Fig. 3. Graph of estimated regression function *PrF* (continuous line) and Monte Carlo estimates for *PrF* (dots) when $K \in [1, 30]$, and $m = 10\,000$. Parameter *T* was chosen at random from interval [50, 250]

Function *PF* appearing in the above formula is of the form:

$$PF(K) = \beta_0 + \beta_1 / K + \beta_2 K + \beta_3 K^2$$

with the following *LS* estimates for unknown coefficient β_i:

$$b_0 = -1.2932,\ b_1 = 1.4272,\ b_2 = -0.05456,\ b_3 = -0.0101$$

The relative prediction error in this case (computed on the basis of the validation set) amounts to 0.00776. The regression model performs really well.

Final remarks

The Monte Carlo experiments - as all computer simulations - are subject to a similar weakness; the results may depend on the specific experiment design. Thus, we propose here to combine the simulation with regression analysis to generalize the results for an arbitrary set of possible design parameters. With the help of the proposed approach we develop models for the value of an unbound blackjack type optimal stopping problems with a linear payoff and exponential step as well as for the risk characteristics of the OSR. The models allow the decision maker to study the risk characteristics of the OSR for a wide range of design parameters. The estimated prediction errors appear to be very small, which indicates that the approach results in the analytical models which are very good approximations of the true relationship.

References

[1] Grzybowski A.Z., Optimal Stopping Rules for Some Blackjack Type Problem, Current Themes in Engineering Science 2009, ed. A. Korsunsky, Melville, American Institute of Physics, New York 2010, 91-100.

[2] Grzybowski A.Z., Monte Carlo Analysis of Risk Measures for Blackjack Type Optimal Stopping Problems, Engineering Letters 2011, 19, 3,147-154 (online version available: 24 August 2011).

[3] Chow Y.S., Robbins H., Siegmund D., Great Expectations: The Theory of Optimal Stopping, Houghton Mifflin, Boston 1971.

[4] Shiryayev A.N., Optimal Stopping Rules, Springer-Verlag, New York 1978.

[5] Bearden J.N., Murphy R.O., Rapoport A., A multi-attribute extension of the secretary problem: Theory and experiments, Journal of Mathematical Psychology 2005, 49, 5, 410-422.

[6] Ferguson T.S., Klass M.J., House-hunting without second moments, Sequential Analysis 2010, 29, 3, 236-244.

[7] Ghasemi A., Yacout S., Ouali M., Optimal stategies for non-costly and costly observations in condition based maintenance, IAENG International Journal of Applied Mathematics 2008, 38, 2, 99-107.

[8] Mihalyi D., Novitzka V., Coalgebras as intrusion detection system, Acta Polytechnica Hungarica 2010, 7, 2, 71-79.

[9] Beibel M., Lerche H.R., A new look at optimal stopping problems related to mathematical finance, Statistica Sinica 1997, 7, 93-108.

[10] Jönsson H., Kukush A.G., Silvestrov D.S., Threshold structure of optimal stopping domains for American type options, Theory of Stochastic Processes 2002, 8, 24,169-177.

[11] Mordecki E., Optimal stopping and perpetual options for Lévy processes, Finance Stoch. 2002, 6, 473-493.

[12] Surya B.A., An approach for solving perpetual optimal stopping problems driven by Lévy processes, Stochastics 2007, 79, 337-361.

[13] Ludkovski M., A Simulation approach to optimal stopping under partial information, Stochastic Processes and Applications 2009, 119, 12, 2071-2087.

[14] Padmanabhan D., Agarwal H., Renaud J.E., Batill S.M., A study using Monte Carlo Simulation for failure probability calculation in reliability-based optimization, Optim. Eng. 2006, 7, 297-316.

[15] Tian X., Benkrid K., Gu X., High performance Monte-Carlo based option pricing on FPGAs, IAENG Engineering Letters 2008, 16, 3, 434-442.

[16] Christensen S., Irle A., Novikov A., An elementary approach to optimal stopping problems for AR(1) sequences, Sequential Analysis 2011, 30, 1, 79-93.

[17] Novikov A.A., Shiryaev A.N., On an effective solution of the optimal stopping problem for random walks, Theory Prob. Appl. 2006, 49, 344-354.

[18] Novikov A.A., Shiryaev A.N., On a solution of the optimal stopping problem for processes with independent increments, Stochastics 2007, 79, 393-406.

[19] Villeneuve S., On threshold strategies and the smooth-fit principle for optimal stopping problems, J. Appl. Probab. 2007, 44, 1, 181-198.