

AN APPROACH TO CUSTOMER BEHAVIOR MODELING USING MARKOV DECISION PROCESS

Ondřej Grunt, Jan Plucar, Markéta Štáková, Tomáš Janečko, Ivan Zelinka

VSB - Technical University of Ostrava
Department of Computer Science
17. listopadu 15, Ostrava-Poruba
Czech Republic

ondrej.grunt@vsb.cz, jan.plucar@vsb.cz, marketa.stakova.st@vsb.cz, tomas.janecko@vsb.cz,
ivan.zelinka@vsb.cz

Abstract: This paper presents an application of Markov Decision Process method for modeling of selected marketing processes. Based on available realistic data, MDP model is constructed. Customer behavior is represented by a set of states of the model with assigned rewards corresponding to the expected return value. Outcoming arcs then represent actions available to the customer in current state. Favourable outcome rate of available actions is then analyzed, with emphasis on suitability of the model for future predictions of customer behavior.

Keywords: markov decision process; marketing; models; probability;

1 Introduction

Since financial institutions such as the one in our article collect extreme amounts of data every day, there is need to sort these bulks of information and extract valuable contexts. These emerging links help marketers to understand sequences of past actions, to learn from them and to extrapolate for future marketing actions. Thus we may include hereby presented methods as a part of strategic marketing planning. Although the area of marketing big data analysis and processing is dynamically developing, it still cannot be considered as prevalent in the daily marketing practice.

Careful observation of customer's actions in the course of time gives us overview of his needs and habits. This basic awareness gives us good starting position for developing meaningful CRM (customer relationship management) through means of premeditated and valid marketing communication. As Wessling [1] states, CRM is mainly shaped by human potential, market processes, technologies and data. It revolves in cycles consisting of:

1. understanding customer's values
2. making contact
3. developing the company-customer relation
4. consolidation
5. decline
6. termination

In majority of above mentioned stages it is necessary to distribute clients among segments by grouping them according to their behavior for the analysis to yield applicable results. Since this strategy is done on historical data, it results in predictions. Traditionally, the businesses would perform SWOT (strengths, weaknesses, opportunities, threats) analysis on manually assembled historical data. However, even though our approach is more demanding, it can bring much deeper and comprehensive insight hence more reliable and clear view.

Predictions can be then used for marketing planning. The aim is to please business's needs by satisfying the client. The knowledge of the market is dependent on sufficient amount of reliable data thus we suggest paying attention to their correct collection. Marketing action should be shaped by several prerequisites to reach strategically valid goal for the enterprise. The communication shall respect the profile of the targeted person. This way, he will obtain an offer for a product he is the most likely interested in. Moreover, we can set the timing of the offer and frequency it might repeat with. In case of having quality input data we are also able to assess which channel to use to approach him.

Some methods such as Markov Chains (MC) thus bring an opportunity to make the communication more cost effective and efficient in order not to overload the client with irrelevant information. This might potentially result in higher awareness and responsiveness to the offer from the part of the customer. In [2], resource allocation is optimized by use of MC model. [3] shows application of MC to show switches between economic expansions and recessions, lately improved in [4]. Customer relationship can be represented by MC as well [5]. However, MC approach may not yield desirable outputs, e.g. efficiency of product offers and campaigns. This led to development of extension to MC called Markov Decision Processes (MDP). Since then, MDP was used for optimization of marketing planning and budgeting [6, 7, 8, 9].

In this paper, MDP based approach to modeling of customer behavior based on data obtained by several financial institutions is presented. Customer behavior is represented by MDP model with states determined by a configuration of three selected attributes of the customer. These attributes were considered to influence the most the choice of action by the customer. The aim of the research was then to evaluate each state correspondingly to initial evaluation and possible outcome of actions taken in that state. From the results, the most profitable (i.e. rewarding) sequence of actions may be identified, providing the institution with insight which policies should be adjusted to increase the return value of the customer.

2 Experiment Design

2.1 Preprocessing Data

Due to complexity and large volume of obtained real data, it was necessary to carry out the preprocessing phase. During this phase, data was cleaned and normalized, reducing complexity of the data set.

Main operations of the preprocessing phase were as follows:

- selection of the most influential attributes
- selection of appropriate records
- handling of incomplete data columns

For the preprocessing, Apache Spark Framework was chosen as a suitable tool. Apache Spark [10] is fast and general-purpose cluster computing system.

Originally, obtained data contained 3.3 billion atomic cells in total, each representing a specific value. After all operations were performed, size of the data set was reduced to 2.1 billion cells, resulting in 36.65% reduction mainly due to sparsity of some data columns.

2.2 Markov Decision Process

Markov Decision Process (MDP) is a discrete time stochastic control process [11, 12]. In situations, where the outcome can be considered partly random and partly under control of some decision maker, MDP can provide mathematical framework for modeling of related decision making. At each time step t , the process is in some state s . Then, the decision chooses one action a available at state s . At the next time step $t + 1$, the process randomly moves into new state s' . Then, decision maker receives reward $R_a(s, s')$ corresponding with s' . State transition function $P_a(s, s')$ represents probability of the process moving into state s' by decision maker choosing action a . In this way, the next state s' depends only on the current state s and action a (thus making it independent of all previous states and chosen actions), therefore satisfying the Markov property.

MDP can be defined as follows:

Definition. Let

- S be a set of states
- A be a set of actions
- $T : S \times A \times S \times \{0, 1, \dots, H\} \rightarrow \langle 0, 1 \rangle$, $T_t(s, a, s') = P(s_{t+1} = s' | s_t = s, a_t = a)$
- $R : S \times A \times S \times \{0, 1, \dots, H\} \rightarrow \mathfrak{R}$, $R_t(s, a, s') = \text{reward for } (s_{t+1} = s', s_t = s, a_t = a)$
- H be horizon over which the customer will act

The goal of MDP can then be summarized as

- Find $\pi : S \times \{0, 1, \dots, H\} \rightarrow A$ that maximizes expected sum of rewards, i.e.

$$\pi^* = \arg \max_{\pi} E \left[\sum_{t=0}^H R_t(S_t, A_t, S_{t+1}) | \pi \right] \quad (1)$$

2.3 State Vector

As previously mentioned, states in the MDP model were determined by a configuration of state vector. Elements of this vector consisted of three attributes from provided data which were considered the most influential on behavior of the customer:

- Monthly change in Balance
- Offer success ratio
- Monthly change in Debt to income (DTI) ratio

Monthly difference was chosen as the best indication of change in customer behavior. Values of elements were then divided into three categories for each element, following comparison to a set of thresholds (categories for each element are shown in Tab. 1).

Table 1: State vector element categories

Attribute	Categories
Balance	Decreasing, Stable, Increasing
Offers	Poor, Average, Good
DTI	Decreasing, Stable, Increasing

Each possible combination of shown categories then resulted in a state. In our research, 27 possible states of the customer were identified.

2.4 MDP Model

Based on the available data, all states were considered to be adjacent to every other state. Actions were represented by a change of category for one or more elements of the state vector (e.g. balance change from stable to decreasing over a month). Directed arcs represent possible movement from state into another states of the model. While this could in theory make the model difficult to be graphically represented and managed, in reality many of these arcs would corresponding to zero probability of movement between states. This is due to the use of input data in the form of monthly differences, which prevent rapid changes in behavior of the customer. Therefore, reasonably lucid MDP model could be obtained (see Fig. 1).

The model was constructed using modified adjacency matrix, which contained outgoing probability of state moving into another state as well as probability of remaining in state. Thickness of arcs in Fig. 1 corresponds to these probabilities. In addition, probabilities of a state being the output of an arc was computed and also visualized in Fig. 1 (light coloration represents states that are rarely an output state, while darker colors represent frequent output states). It could be seen, that both thickness of arcs and coloration point to state 11 as the one which most of the customers reached and remained in (i.e. making state 11 absorbing state in this model). This can be explained by state 11 representing customer with no significant monthly changes in both balance and DTI and with poor offer success rates. Simply put, we can consider such customer to have either stable or seldomly used account.

Another important observation in Fig. 1 is inactivity of two thirds of the states present in the model. This is mainly due to limited data on offer campaigns and their success rates. It should be also noted, that this is the main reason for prevalence of states representing customers with poor or no response to offers made by institution.

Next step of our research was to compute reward function values for each of the active states in MDP model. For this, value iteration method was chosen as the suitable tool. However, some of the states can be considered undesirable for the institution, such as states corresponding to customer with decreasing balance and/or increasing DTI. Similarly, beneficial states were identified as well. Then, initial rewards were assigned to these states ranging from -10 and -5 for undesirable ones to 5 and 10 for desirable ones. Remaining states were assigned value of 0. Resulting MDP model is shown in Fig. 2, where green states represent desirable states and red states represent undesirable ones, with exact set of reward values also presented in Tab. 2 and Tab. 3:

Table 2: Reward values before value iteration

State	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Value	-10	-5	0	-5	0	5	0	0	5	-10	0	0	-5	0

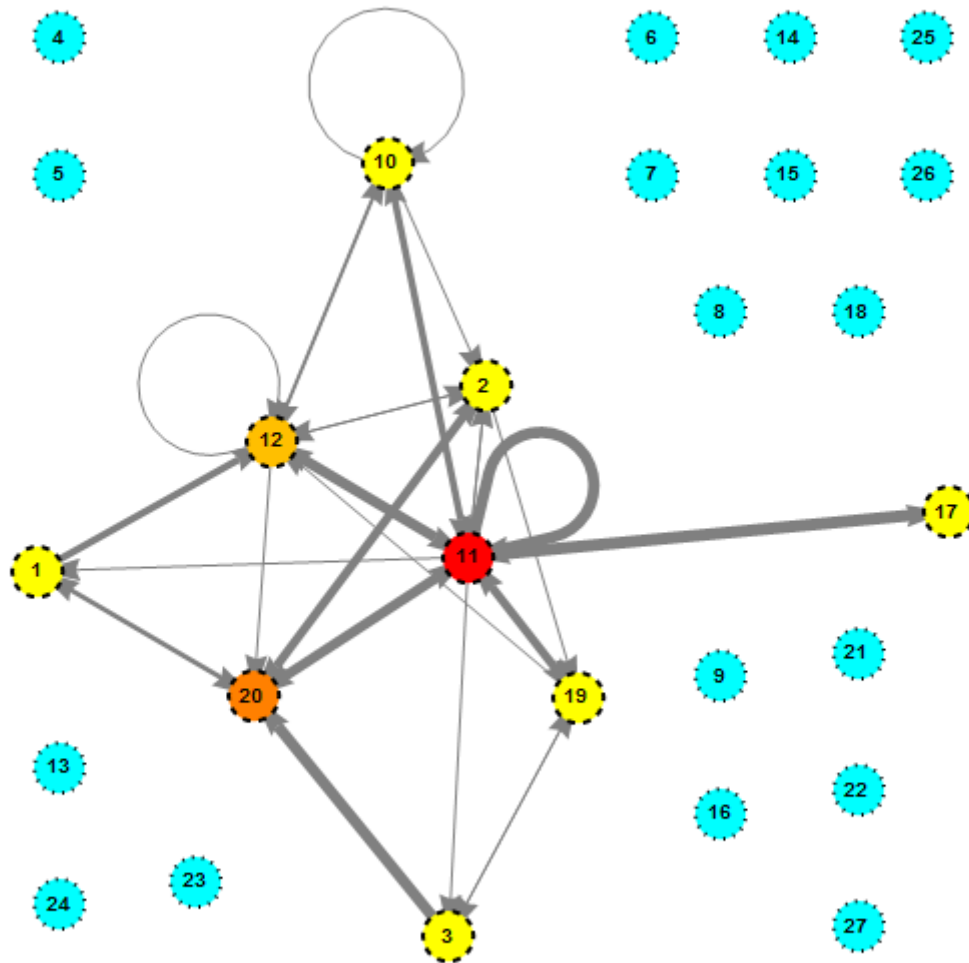


Figure 1: MDP model.

Table 3: Reward values before value iteration

State	15	16	17	18	19	20	21	22	23	24	25	26	27
Value	5	0	0	10	-10	0	0	-5	0	5	0	5	10

First observation is the lack of positively rewarded states amongst the active ones. Therefore, it could be expected that there will be no positively rewarded state following value iteration.

3 Results

First results obtained by MDP model will be discussed in this section. Following the observation of zero presence of positively rewarded states in the MDP model, adjustments to the model and data mining will be discussed as well.

3.1 MDP Outputs

Convergence was reached after 120 iterations resulting in MDP model shown in Fig. 3. Computed reward values are shown in Tab. 4 and Tab. 5 and are represented in Fig. 3 by varying degrees of red coloration of

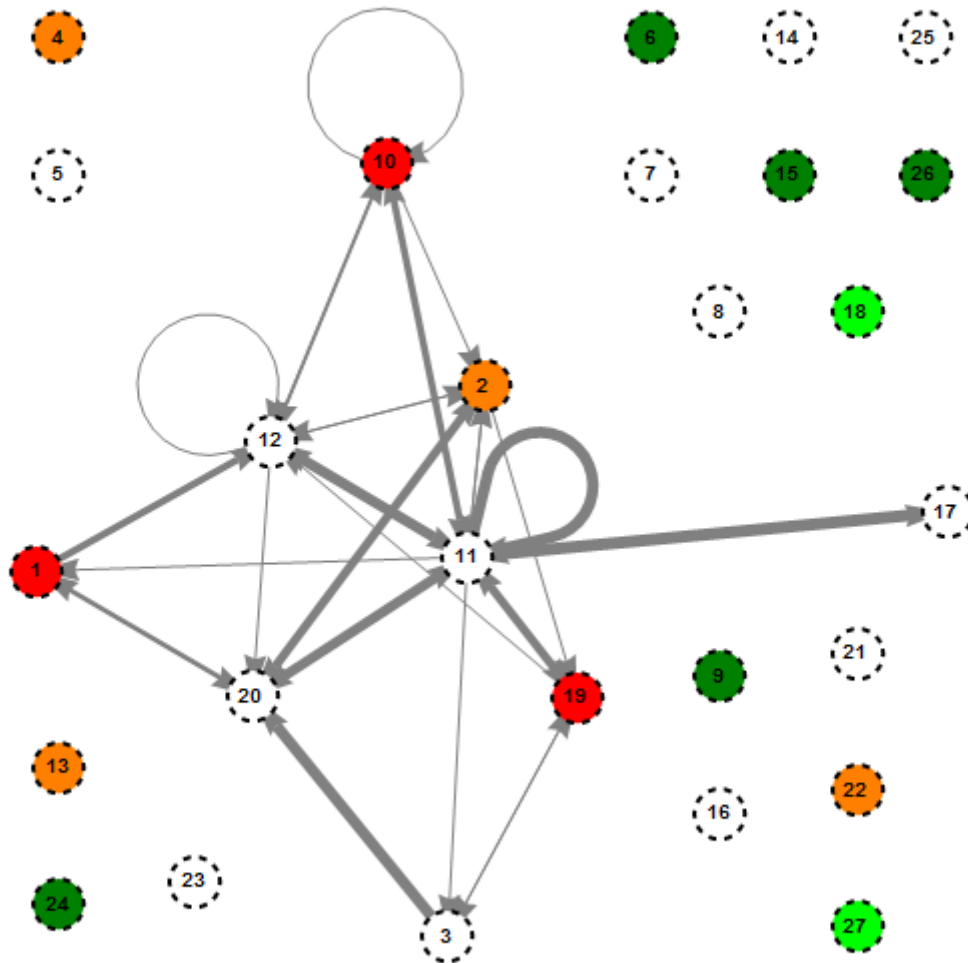


Figure 2: MDP model with initial reward values.

active states.

Table 4: Reward values following value iteration

State	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Value	-12.9992	-8.6536	-3.5338	-5	0	5	0	0	5	-13.8883	-2.6871	-3.3500	-5	0

Table 5: Reward values following value iteration

State	15	16	17	18	19	20	21	22	23	24	25	26	27
Value	5	0	-2.4184	10	-12.6125	-3.3060	0	-5	0	5	0	5	10

As it was predicted, only initially negatively rewarded states had influence on reward values of active states (i.e. all formerly "neutral" states were influenced by states 1, 2, 10 and 19).

Normally, this step in MPD model evaluation would be followed by identification of optimal policy for the institution to maximize the return value of the customer. However, as it was previously mentioned, no positively

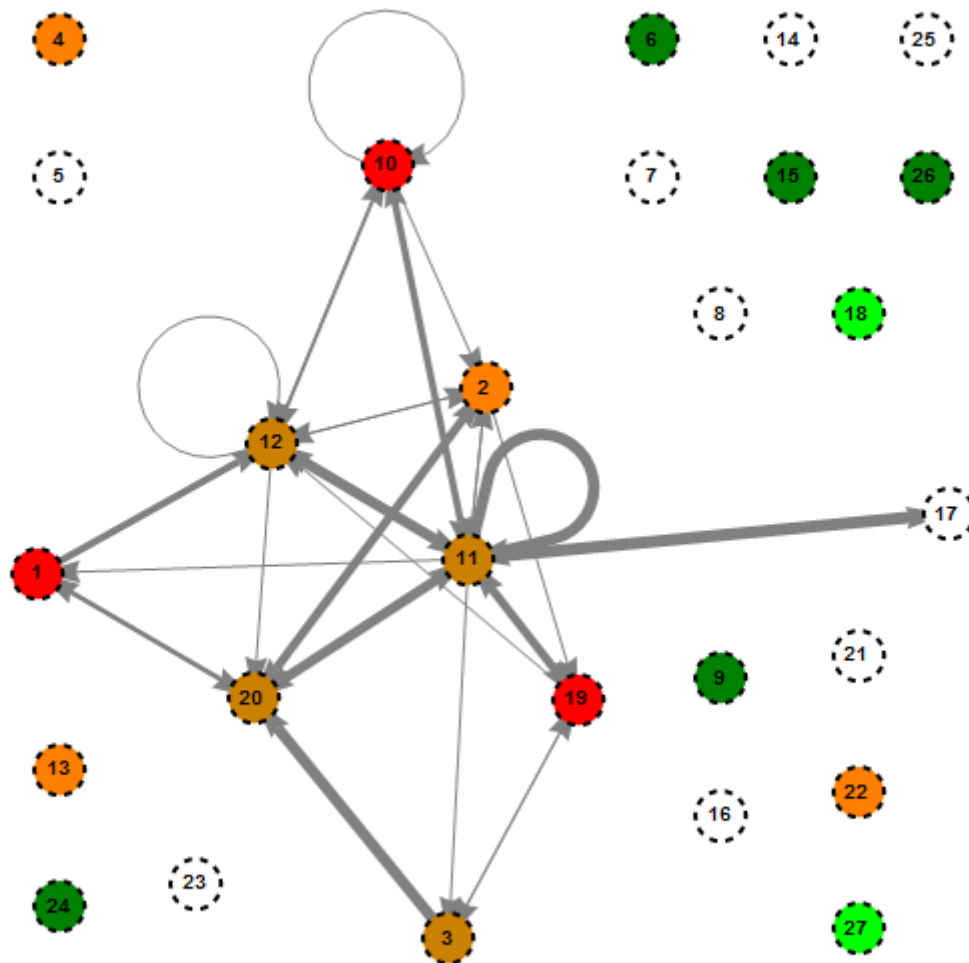


Figure 3: MDP model with updated reward values.

rewarded active state is present in the MDP model. Therefore, in our research, the optimal policy would be for the customer to reach state 11 and remain in that state to minimize total loss (contrary to the intended maximization of customer return value). Again, as state 11 represents customer with no significant monthly changes in both balance and DTI and with poor offer success rate, it could not be considered optimal policy from view of the institution.

3.2 Future Research

The lack of positively rewarded states in the model can be seen as its biggest drawback. This was mainly due to sparsity of obtained data related to offer success rates and partially DTI. Therefore, more complete additional data have to be obtained and used as a basis for future model.

Another complication was the chosen nature of the state vector. While all selected elements of the vector are considered to be important to determine return value of the customer, introduction of more elements into the vector is in order. In this way, the MDP model should be less dependent on some variables. However, this may introduce complication in form of significantly more configurations of the state vector, resulting in model with too many states to be manageable. Because of this, these configurations will be combined into several segments, representing states in resulting MDP model.

Our next proposed step is to use evolutionary algorithms (EA) in two major ways:

- The first one is to use "classical" EA to estimate parameters of the MDP model to fit the behavior of the model more closely to real data. This should be a simple task, as the MDP model structure will be designed in accordance with demands of the institution and EA will then be used for estimation of numerical parameters of the MDP model. For those purposes, one of algorithms like differential evolution [13], SOMA [14] or particle swarm [15] will be used.
- The second one is to use symbolic regression (SR) like genetic programming [16], grammatical evolution [17] or analytic programming [18]. In this part, SR will be used to synthesize MDP models including its parameters according to used datasets. This approach has already been used in our different research focused on Petri nets synthesis.

4 Conclusion

In this paper, MDP are used for modeling of selected marketing process to maximize return value of the customer. MDP model using states in form of configuration of state vector was constructed, using modified adjacency matrix derived from provided realistic data. Then, reward values were computed for each state following value iteration method application.

Original aim of presented work was to identify the best course of actions forming optimal policy for the institution to increase the return value of its customers in its future campaigns. However, due to sparse nature of available realistic data combined with the choice of state vector, two thirds of states in MDP model were inactive, since no customer reached them. As all of states which are considered to be desirable from the view of the institution were part of that group, the active states either remained neutral or were assigned negative reward following value iteration method. Therefore, instead of maximizing potential reward, the model could only be used to minimize potential loss. This policy was then represented by customer reaching and remaining in absorbing state of the MDP model.

As the performed research did not avail desirable outcome, further adjustments to MDP model construction and data gathering were discussed and are currently being applied. In future works, more attributes influencing the customer behavior will be introduced into MDP model creation by use of segmentation instead of states represented by state vector configuration. This should allow the model to be less dependent on the nature of single variable while allowing some degree of control over the number of states in MDP model. Better data mining and data processing methods will be applied as well. Also, evolutionary algorithms will be used for achieving better fitness of the MDP model to given data for more accurate prediction of the customer behavior.

Acknowledgement: The following grants are acknowledged for the financial support provided for this research: Grant Agency of the Czech Republic - GACR P103/15/06700S, Grant of SGS No. SGS 2017/134, VSB-Technical University of Ostrava. The Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project IT4Innovations excellence in science - LQ1602. Research presented in this article was conducted in collaboration with PricewaterhouseCoopers Česká republika, s.r.o.

References

- [1] Wessling, H.: Aktivní vztah k zákazníkům pomocí CRM: strategie, praktické příklady a scénáře. Praha: Grada, 2003. Manažer. ISBN 8024705699.
- [2] DeMaagd, N.: The Application of Stochastic Processes in the Sciences: Using Markov Chains to Estimate Optimal Resource Allocation, Population Distribution, and Gene Inheritance. Honors Projects. Paper 290 (2014).
- [3] Hamilton, J.D: A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57(2), 357–384 (1989).
- [4] Calvet, L.E.: How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes. *Journal of Financial Econometrics* 2 (1), 49–83 (2004).
- [5] Pfeifer, P.E., Carraway, R.: Modeling Customer Relationships as Markov Chains. *Journal of Interactive Marketing* 14(2), 43–55 (2000).
- [6] Labbi, A., Berrospi, C.: Optimizing marketing planning and budgeting using Markov decision processes: an airline case study. *IBM Journal of Research and Development - Business optimization* 51(3), 421–431 (2007).
- [7] Bäuerle, N., Rieder, U.: MDP algorithms for portfolio optimization problems in pure jump markets. *Finance and Stochastics* 13(4), 591–611 (2009).
- [8] Perez, I., Hodge, D., Le, H.: Markov decision process algorithms for wealth allocation problems with defaultable bonds. *Advances in Applied Probability* 48(2), 392–405 (2016).

- [9] Rösch, A., Schmidbauer, H.: Action Selection in Customer Value Optimization: An Approach Based on Covariate-Dependent Markov Decision Processes. Proceedings of the 2009 International Conference on Data Mining DMIN 09, Las Vegas. ISBN: 1-60132-099-X.
- [10] Zaharia, M.: Apache Software Foundation, UC Berkeley AMPLab, Databricks. <http://spark.apache.org/docs/latest/> (2017). [Online; accessed 3-May-2017]
- [11] Bellman, R.E.: A Markovian Decision Process. *Journal of Mathematics and Mechanics* 6(5), 679–684 (1957).
- [12] Howard, R.A.: *Dynamic programming and Markov processes*. Cambridge, Massachusetts, 1960. The M.I.T. Press.
- [13] Storn, R., Price, K.: Differential evolution: a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization* 11(4), 341–359 (1997).
- [14] Davendra, D., Zelinka, I.: *Self-Organizing Migrating Algorithm: Methodology and Implementation*. Springer, 2016. ISBN 978-3-319-28161-2.
- [15] Clerc, M.: Particle swarm optimization. Vol. 93. John Wiley & Sons, 2010. Kennedy, James. "Particle swarm optimization." In *Encyclopedia of machine learning*, pp. 760-766. Springer US, 2011.
- [16] Koza, J. R.: *Genetic programming: on the programming of computers by means of natural selection*. Vol. 1. MIT press, 1992.
- [17] O'Neil, M., Ryan, C.: Grammatical evolution. In *Grammatical Evolution*, pp. 33-47. Springer US, 2003. Harvard
- [18] Zelinka, I., Davendra, D., Senkerik, R., Jasek, R., Oplatkova, Z.: Analytical Programming - a Novel Approach for Evolutionary Synthesis of Symbolic Structures, *Evolutionary Algorithms*, Prof. Eisuke Kita (Ed.), InTech, 2011. DOI: 10.5772/16166. Available from: <https://www.intechopen.com/books/evolutionary-algorithms/analytical-programming-a-novel-approach-for-evolutionary-synthesis-of-symbolic-structures>