Scientific Journals
Maritime University of Szczecin

Zeszyty Naukowe
Akademia Morska w Szczecinie

2014, 38(110) pp. 69–74
ISSN 1733-8670

2014, 38(110) s. 69–74

The problem of the special technical system operator suitability preliminary assessment

Marek Malarski

Warsaw University of Technology, Faculty of Transport
00-662 Warszawa, ul. Koszykowa 75, e-mail: mma@ wt.pw.edu.pl

Key words: special operator, ATC controller, rough sets theory, training, recruitment

Abstract

The main aim of this paper is to introduce the possibility for applying rough sets theory to analyze the test result which candidates for special operator (ATC controller) must took while their recruitment. Recruitment after reaching the person being trained, consisting of two parts (theoretical and practical) and lasting more than two years. The candidate must be distinguished features such as an excellent orientation in space, perception, logical thinking, divided attention, stress resistance, ability to plan, very good health and a very good knowledge of Polish and English. Paper consists a detailed description of the recruitment for special operator – ATC controller position and brief introduction to the rough sets theory including basic concepts and methodology.

The analytical part of this paper describing set of psychophysical characteristics of operator, means by which and based on actual data, decision table is created. Afterwards all data are analyzed in the special computer software. The paper ends with a summary, the results of the analysis are discussed.

Introduction

Jobs operators many special technical systems is very responsible and often from them depends the safety of people. At the stage of the recruitment of these operators are subject to many, often very sophisticated tests evaluating their suitability for the future of the profession. After the recruitment of operators are subject to the training specialist. This training is often lengthy and very costly. Thus, in many areas of economic life are attempts to improve the initial evaluation of the suitability of the candidate for the operator. One possibility here is to use data mining methods and knowledge discovery. In the modern world there is a great demand for such methods. They allow you to systematize a large amount of data and pull out on the basis of their conclusions, or make a decision. One of the tools of data mining is the method of rough sets used in this article to analyze the results of the qualification tests of the technical system operators – ATC controllers.

An example of the recruitment process for a course of special technical system operator

Sample recruitment rate highly specialized provider of special technical system is carried out by a specialized unit in the cycle once or several times a year [1]. The very process of recruitment is very long and optimized to determine the best candidates. An example of the recruitment process consists of five steps (Fig. 1).

Pre-qualification at stage 1 analysis submitted by the candidate application, which also includes a preliminary check of the English language. In stage 2, candidate approaches the two written tests:

- aptitude test is mainly to verify perception, reasoning skills, spatial imagination candidate,
- "Placement Test" – verifies the correctness of grammar and vocabulary of English Candidate.

It is essential to assess the suitability of the candidate to test stages 3 and 4. Stage 3 is the same for the whole of Europe aptitude test computer [2]:

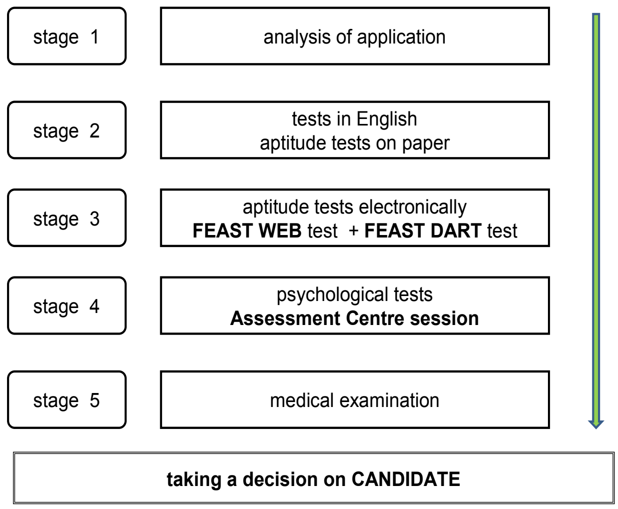

Fig. 1. Sample selection stages Candidates for exchange operator specific technical system

- **FEAST** (*First European ATC Controller Selection Test*) **WEB** – group 6 of tests:
  - the ability to make decisions;
  - logical reasoning;
  - visual perception;
  - multi-tasking;
  - spatial orientation;
  - the ability to maintain continuous concentration;
  - listening comprehension in English;
- **FEAST DART** (*Dynamic ATC Radar Test*) – checks multitasking, accuracy and efficiency of the candidate on 6 simulator tests.

The result is a rating scale [0, 9] 11 candidate competences. This is complemented by step 4 includes additional psychological tests and so. Session Assessment Centre, as evaluated by a group of experienced assessors. Features especially evaluated at this stage are:

- ability to work in a team;
- the ability to make their own decisions and bear the consequences;
- communication skills;
- ability to cope with stress;
- easy to adapt to the situation;
- the rate of acquisition of new skills;
- ability to adapt to existing procedures.

Recruitment at this level close personal interview conducted in English and Polish.

The last step (stage 5) is recruiting medical examination to verify the health of the candidate in terms of contraindications to perform the duties of the operator. The negative outcome of these studies will disqualify the candidate.

Further recruitment path includes training (often very expensive). In the analyzed example are:

- theoretical basic course (BASIC TRING) – about 16 weeks;
- the course of the powers (RATING TRING) – about 8 weeks;
- training simulator – about three months;
- practice in the workplace – 1–2 years.

Each step ends eligible for further examinations parts. The training process is very costly. The training operator thus depends on the initial qualification that the percentage of candidates successfully completing the course the powers, was the maximum.

Selected elements of the rough sets theory used in the analysis

Information System [inter alia 3, 4] is an ordered:

$$SI = \{\boldsymbol{U}, \boldsymbol{A}, \boldsymbol{V}, f\} \tag{1}$$

where:

$\boldsymbol{U}$ – non-empty, finite set of objects – called the universe;

$\boldsymbol{A}$ – non-empty, finite set of attributes;

$\boldsymbol{V} = \bigcup_{a \in A} V_a$, where: V_a – the attribute field $a \in \boldsymbol{A}$;

$f: \boldsymbol{U} \times \boldsymbol{A} \to \boldsymbol{V}$ – information function, such that $\forall_{u \in \boldsymbol{U}}$ and $\forall_{a \in \boldsymbol{A}} : f(a,u) \in V_a$.

A special case *SI* is a ***Decision Table*** *TD*. It is created by the award disjoint sets of conditional attributes $\boldsymbol{C}$ and decision attributes $\boldsymbol{D}$ such that $\boldsymbol{A} = \boldsymbol{C} \cup \boldsymbol{D}$. *TD* advantage is the ability to present a set of conditions that must be met in order to be executed corresponding set of decisions

$$TD = \{\boldsymbol{U}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{V}, f\} \tag{2}$$

where:

$\boldsymbol{U}$ – non-empty, finite set of objects;

$\boldsymbol{C}$ – non-empty, finite set of conditional attributes;

$\boldsymbol{D}$ – non-empty, finite set of decision-making attributes;

$\boldsymbol{V} = \bigcup_{a \in \boldsymbol{C} \cup \boldsymbol{D}} V_a$, where V_a – the attribute field $a \in \boldsymbol{C} \cup \boldsymbol{D}$;

$f: \boldsymbol{U} \times (\boldsymbol{C} \cup \boldsymbol{D}) \to \boldsymbol{V}$ – information function, such that $\forall_{u \in \boldsymbol{U}}$ and $\forall_{a \in \boldsymbol{C} \cup \boldsymbol{D}} : f(a,u) \in V_a$.

For *TD* introduced is the concept of ***rule-making*** – this is the form of conditional sentences, in which each object $u \in \boldsymbol{U}$. *Rule of decision-making* so called function: $g : (\boldsymbol{C} \cup \boldsymbol{D}) \to \boldsymbol{V}$ if there exist $x \in \boldsymbol{U}$ such that $g = f_x$.

Decision rules can be divided into two groups:

- decision rule $g = f_x$ in *TD* is **deterministic** when:

$$\underset{\substack{x,y\in \boldsymbol{U} \\ x\neq y}}{\forall} \cdot f_x|\boldsymbol{C} = f_y|\boldsymbol{C} \Rightarrow f_x|\boldsymbol{D} = f_y|\boldsymbol{D} \tag{3}$$

– otherwise it is **non-deterministic** rule.

TD is ***well defined*** $\Leftrightarrow$ all decision rules are deterministic.

TD is ***poorly defined*** $\Leftrightarrow$ in the *TD* exist non-deterministic decision rule.

Let $x, y \in \boldsymbol{U}$ and $\boldsymbol{B} \subseteq \boldsymbol{A}$:

- x, y are ***identifiable*** by $\boldsymbol{B}$, if there exists such $a \in \boldsymbol{B}$, that $f(a, x) \neq f(a, y)$;
- x, y are ***indistinguishable*** by $\boldsymbol{B}$, if there exists such $a \in \boldsymbol{B}$, that $f(a, x) = f(a, y)$.

With every subset of attributes $\boldsymbol{B} \subseteq \boldsymbol{A}$ linked to a binary relation $I(\boldsymbol{B})$ called the relation of ***indiscernibility***

$$I(\boldsymbol{B}) = \{(x, y) \in \boldsymbol{U} \times \boldsymbol{U} : \underset{a\in \boldsymbol{B}}{\forall} f(a, x) = f(a, b)\} \tag{4}$$

If $(x, y) \in I(\boldsymbol{B})$ than the objects x and y are indistinguishable due to a subset of attributes $\boldsymbol{B}$.

For each relationship *indiscernibility* we can distinguish disjoint subsets called ***classes of abstraction*** / ***indiscernibility*** (or ***sets of elementary***) of this relationship, which divides it set in which it is defined. The classes are called ***B-elementary*** sets. The symbol $B(x)$ denote the relation ***equivalence class*** $I(\boldsymbol{B})$ containing the object x.

A special case of *indiscernibility relations* is the *relation indiscernibility with respect to the decision*. In this relation objects x i y are *indistinguishable* if they have the same decision-making values (conditional attributes values may differ).

May $TD = \{\boldsymbol{U}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{V}, f\}$ and $\boldsymbol{B} \subseteq \boldsymbol{C}$. Indiscernibility relation with respect to decision $\boldsymbol{D}$ conditioned by a set of attributes $\boldsymbol{B}$ express

$$I(\boldsymbol{B}, \boldsymbol{D}) = \\ = \{(x, y) \in \boldsymbol{U} \times \boldsymbol{U} : (x, y) \in I(\boldsymbol{B}) \vee f(x, d) = f(y, d)\} \tag{5}$$

Each set of objects $\boldsymbol{X} \subseteq \boldsymbol{U}$ can be described by a set of attributes $\boldsymbol{B} \subseteq \boldsymbol{A}$ in a precise or approximate.

May $SI = \{\boldsymbol{U}, \boldsymbol{A}, \boldsymbol{V}, f\}$ and $\boldsymbol{B} \subseteq \boldsymbol{A}$. The set $\boldsymbol{X} \subseteq \boldsymbol{U}$ is a set ***B-exact-*(*B-definable*)** if it is a finite union of the sets of ***B-elementary***. Each set, which is not the sum of ***B-elementary*** sets is called a ***B-approximate***.

Rough sets are divided into classes. We are able to define the four of them.

a) Class ***approximately B-definable*** sets:

$$B_*(X) \neq \varnothing \wedge B^*(X) \neq U \tag{6}$$

The set $\boldsymbol{X}$ is a set of *approximately B-definable* when using the set $\boldsymbol{B}$ we are able to determine whether some elements of $\boldsymbol{U}$ belong to $\boldsymbol{X}$ or $\neg\boldsymbol{X}$.

b) Class ***internally B-undefined*** sets:

$$B_*(X) \neq \varnothing \wedge B^*(X) \neq U \tag{7}$$

The set $\boldsymbol{X}$ is a set of *internally B-undefined* when using the set $\boldsymbol{B}$ we are able to determine whether some elements of $\boldsymbol{U}$ belong to $\neg\boldsymbol{X}$ while we can not clearly determine whether any element belongs or does not belong to $\boldsymbol{X}$.

c) Class ***externally B-undefined*** sets:

$$B_*(X) \neq \varnothing \wedge B^*(X) = U \tag{8}$$

The set $\boldsymbol{X}$ is a set of *externally B-undefined* when using the set $\boldsymbol{B}$ we are able to determine whether some elements of $\boldsymbol{U}$ belong to $\boldsymbol{X}$ but we can not clearly determine whether any element belongs or does not belong to $\neg\boldsymbol{X}$.

d) Class ***completely B-undefined*** sets:

$$B_*(X) = \varnothing \wedge B^*(X) = U \tag{9}$$

The set $\boldsymbol{X}$ is a set of *completely B-undefined* when using the set $\boldsymbol{B}$ we are not in a position to decide whether certain elements of $\boldsymbol{U}$ belong to $\boldsymbol{X}$ or $\neg\boldsymbol{X}$.

One of the main problems that we encounter when analyzing (particularly extensive) data, there are excess. But we must remember that sometimes not all attributes are needed in the decision-making process. To find and eliminate redundant data in the *SI* use the concept of ***reduct of attributes***.

Let $SI = \{\boldsymbol{U}, \boldsymbol{A}, \boldsymbol{V}, f\}$ it be an information system and $\boldsymbol{B} \subseteq \boldsymbol{A}$:

- say that attribute a is ***unnecessary in B*** if: $I(\boldsymbol{B}) = I(\boldsymbol{B} - \{a\})$, otherwise the attribute and the attribute is ***indispensable in B***;
- set of attributes $\boldsymbol{B}$ is independent in *SI* if all attributes belonging to it are ***necessary***, otherwise the set $\boldsymbol{B}$ is a set of ***subsidiary***;
- set of attributes $\boldsymbol{G}$, such that $\boldsymbol{G} \subseteq \boldsymbol{B}$ is called the ***set of attributes B reduct*** in the *SI* if the set is independent, and the equality

$$I(\boldsymbol{B}) = I(\boldsymbol{G})$$

Reduct is the smallest subset of attributes that enables the same classification of objects universe as a full set of attributes. We can therefore say that the attributes that do not belong to the reducts are superfluous to the classification of objects (their elimination will not change the overall characteristics of the set).

The core of the attributes $CORE(\boldsymbol{B})$ is the set of all necessary attributes in the set $\boldsymbol{B}$. It is therefore a set of attributes, which is contained in all reducts. Express his dependency

$$CORE(\boldsymbol{B}) = \bigcap_{R \in RED(\boldsymbol{B})} R \tag{10}$$

where: $\boldsymbol{RED}(\boldsymbol{B})$ – the set of all reduct set of attributes $\boldsymbol{B}$ in SI.

It can be concluded that the core attributes is the most important subset of attributes – the removal of any of its elements will result in a complete change in the classification of the elements of the universe.

Let $SI = \{\boldsymbol{U}, \boldsymbol{A}, \boldsymbol{V}, f\}$ it be an information system and $\boldsymbol{C} \subseteq \boldsymbol{A}$ and $\boldsymbol{D} \subseteq \boldsymbol{A}$:

- attribute $a \in \boldsymbol{C}$ is ***D**-redundant* is set of $\boldsymbol{C}$, if $POS_C(\boldsymbol{D}) = POS_{C-\{a\}}(\boldsymbol{D})$ otherwise the attribute a is ***D**-essential*;
- if all attributes $a \in \boldsymbol{C}$ are ***C**-necessary* in the set $\boldsymbol{C}$, the set $\boldsymbol{C}$ is a set of ***D**-independent*;
- subset $\boldsymbol{C}^\circ \subseteq \boldsymbol{C}$ is $\boldsymbol{D}_{\text{-reduct}}$ set $\boldsymbol{C}$, if the set $\boldsymbol{C}^\circ$ is ***D**-independent* and

$$POS_C(\boldsymbol{D}) = POS_{C^\circ}(\boldsymbol{D}) \tag{11}$$

The set of all attributes of ***D**-necessary* in the set $\boldsymbol{C}$ is called ***D**-core* set $\boldsymbol{C}$, we denote $CORE_{\boldsymbol{D}}(\boldsymbol{C})$ and express the dependence

$$CORE_{\boldsymbol{D}}(\boldsymbol{C}) = \bigcap RED_{\boldsymbol{D}}(\boldsymbol{C}) \tag{12}$$

where: $RED_{\boldsymbol{D}}(\boldsymbol{C})$ – the family of all $\boldsymbol{D}_{\text{-reduct}}$ set $\boldsymbol{C}$.

If $SI = \{\boldsymbol{U}, \boldsymbol{A}, \boldsymbol{V}, f\}$ an information system such that $\boldsymbol{U} = \{u_1, u_2, ..., u_n\}$ and $\boldsymbol{A} = \{a_1, a_2, ..., a_n\}$ that ***discernible matrix*** $\boldsymbol{P}(SI)$ for information system to save

$$p_{i,j} = \{a \in A : f(u_i, a) \neq f(u_j, a)\} \text{ for } i, j = 1,...,n \tag{13}$$

Often used instead of discernible matrix is an ***differentiation array*** $\boldsymbol{T}(SI)$ determined dependence

$$t_{i,j}(k) = \begin{cases} 0 - f(u_i, a_k) = f(u_j, a_k) \\ 1 - f(u_i, a_k) \neq f(u_j, a_k) \end{cases} \tag{14}$$

where: $i, j = 1, \dots, n$; $n = |U|$; $j > i$; $k = 1,..,|A|$.

Sample analysis of pre-qualification

For example, analyzed the data contains a sample of the results of 26 candidates who are in one of the calls successfully passed the qualification, and 14 of them after further course obtained a license (Decision YES) [5]. Attributes conditional adopted in the analysis are highlighted in tests FEAST eleven attributes (competency) are presented in table 2.

Table. 2. Competencies evaluated during testing FEAST (own study using [2, 6])

Designator	Competency	Feature / explanation
A1	Dual Processing	dual processing / divided attention
A2	English – listening and comprehension	English – listening and comprehension
A3	Heading and range	short-term memory / ability to quickly assimilate the principles and their use
A4	Matches noted	matches noted / eye-hand coordination, the ability to distinguish shapes, quick reflexes
A5	Mean reaction time	mean reaction time / quick decision-making
A6	Messages obeyed	precision work / correctness of completed tasks
A7	Multi Tasking – Alertness	multitasking, resistance to stress, self-control, divided attention / ability to maintain a high state of concentration in stressful situations
A8	Planning – ATC	ability to plan – prediction / ability to quickly solve algebraic expressions
A9	Single Processing	single processing / precision work
A10	Sorting ability	efficient collection and analysis of information obtained from multiple (different) source / ability to filter incoming information
A11	Visualisation	spatial imagination

Decision table

According to the previously presented definition (2) takes the form of a decision table:

$$TD = \{\boldsymbol{U}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{V}, f\}$$

where:

the universe $\boldsymbol{U}$ – a set of candidates:
$\boldsymbol{U} = \{1, 2, 3, ..., 26\}$;

conditional attributes: $A \in \boldsymbol{A}$, $A \in [0, 9]$:
$\boldsymbol{A}$ = {Dual Processing, English – listening and comprehension, Heading and range, Matches noted, Mean reaction time, Messages obeyed, Multi Tasking – Alertness, Planning – ATC, Single Processing, Sorting ability, Visualisation};

decision attribute $\boldsymbol{D}$ – the final result of recruitment:

$\boldsymbol{D} = \{YES, NO\}$;

$\boldsymbol{V} = V_{A1} \cup V_{A2} \cup V_{A3} \cup V_{A4} \cup V_{A1} \cup V_{A6} \cup \cup V_{A7} \cup V_{A8} \cup V_{A9} \cup V_{A10} \cup V_{A11} \cup V_d$;

$f : \boldsymbol{U} \times (\boldsymbol{C} \cup \boldsymbol{D}) \rightarrow V$ – decision function: $\forall_{x \in U, a \in C \cup D} f(a, x) \in V_a$.

Table 3. An example of a decision table (excerpt)

Candidate	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	*d*
1	6	4	7	5	9	7	8	8	4	5	7	*YES*
2	7	8	9	6	8	8	7	8	7	7	9	*YES*
3	6	5	5	5	3	7	8	8	7	7	6	*NO*
25	6	7	7	6	7	6	6	9	9	7	8	*NO*
26	6	4	6	5	7	8	8	7	8	7	6	*YES*

The information system consists of 26 elementary sets, which are divided between two classes of decision {*YES*, *NO*}. In the class of decision {*YES*} there were 14 objects. In the class of decision {*NO*} there were 12 objects. Decision rules are therefore deterministic (3) and distinguishable (4). There is also no need for approximation sets of attributes (5) – (9). Basic research related to so determining the reduct set – minimal sets of attributes for which there is provided a previous classification of objects (10). In the analyzed example, all existing reductive (57 reducts) based on the concept of differentiation matrix, was determined by ROSE2 [7]. Such calculations were also carried out in [6].

Examples reducts are:

1: {A3, A8, A9, A11},
2: {A1, A2, A5, A10, A11},
…
27: {A4, A9, A11},
…
57: {A1, A2, A5, A7, A10}.

Generated reducts have a length of three to five attributes. Even a cursory analysis indicates their lack of a core set of reducts (11)–(12) – which is the product of a logical set of all reducts. It also has been found [6].

One of the most important stages of the analysis is to determine the decision rules. Various algorithms for induction of decision rules [7] tested include in [6, 8]. Generates standard here as the shortest minimum number of decision rules that cover all cases. With low frequencies set UNIVERSE (26 objects), comparable to the number of conditional attributes (11 attributes), you should expect a relatively low coverage factors and ambiguity selection rules. Interesting results were obtained using the extended rule – making; for example, object > value , the object < value of praying in [7]. Example obtained here 13 decision rules: 6 of them has decided {*YES*}, and another 7 {*NO*} decision. From the obtained 5 rules for conditional uses two attributes, and 8 rules – one attribute.

The rule with the highest degree of coverage is: "If in the course of testing FEAST test score for the 7 or 5 for the ability to Visualization it was licensed ATC controller" – coverage of 42.86%.

In addition, the attributes of the A6, A9 and A11 were considered unsuitable in the making of decision rules. "Useless" are competency: "Messages obeyed", "Single processing", "Sorting ability", which in practice is not true.

In addition, few studies verified the accuracy of data classification. Standard [7] classification is performed by cross-validation (*k-fold cross validation*). It involves the random distribution of the data set for *k* equal parts, and then are learned classifier on one of these parts (training set) and checked on a set of *k* – 1 other (test set). This process is repeated until the each part of the set was used as a training set and test set. In addition, benefited from the validation stratified, because it ensures that allocation of objects to retain balance between the classes of decision-making [6, 8]. The result – the average classification accuracy is the ratio of the number of correctly classified test objects to the total number of test objects including the classification error. This result in the best case, it is – on average 60% of all objects were classified correctly, with an average error of about 28%.

Conclusions

This paper aims to verify the effectiveness of pre-qualification of candidates for operators – ATC controllers of different specializations, carried out the latest modern methods in the EU. Access to the test results and recruitment is very limited. In addition, several years ago, replaced by a test FEAST Strip Display Management, a new version of FEAST Dynamic ATC Radar Test. Moreover, in Poland FEAST test results are supplemented with additional qualification tests. The main objective of recent work under the guidance of the author is to determine the possible relationship between the objective results of the initial qualifications and licensing trainee operator – ATC controller. The analyzes carried out show that the relationships they exist, but require research on a larger sample than a single intake.

The results obtained prove that the analysis of the qualification tests using rough sets may be used to verify the results of the candidates during the tests. Note, however, that the recruitment of the operator – ATC controller seeks to identify people who every day will ensure the safety of Polish air transport. Therefore, the result obtained on the possibility of resignation from 3 "competence" should be assessed very carefully.

References

1. The recruitment process for the course for candidates for the position of an air traffic controller and the elements of the course at the Centre for Aviation Training. PANSA, 2012.
2. First European ATC Controller Selection Test (FEAST). EUROCONTROL, https://www.eurocontrol.int/services/first-european-air-traffic-controller-selection-test-feast
3. MRÓZEK A., PŁONKA L.: Analiza danych metodą zbiorów przybliżonych. Zastosowanie w ekonomii, medycynie i sterowaniu (Analysis of the data by rough sets. Applications in economics, medicine and control). Akademicka Oficyna PLJ, Warszawa 1999.
4. PAWLAK Z.: Zbiory przybliżone. Nowa matematyczna metoda analizy danych (Rough sets. A new mathematical method for analyzing data). http://bcpw.bg.pw.edu.pl/Content/2031/zbiory_przyblizone_nowa_mat._metoda_analizy_danych.pdf,
5. Anonymous calls the results of one of the candidates for the course controller ATC, 2012.
6. PŁUDOWSKA M.: Ocena predyspozycji operatora – kontrolera ruchu lotniczego (Rating predisposition operator – an air traffic controller). Job Eng. under the direction of M. Malarski, WT PW, 2014.
7. Zakład Inteligentnych Systemów Wspomagania Decyzji PPoz., User's Guide ROSE2, http://idss.cs.put.poznan.pl/site/fileadmin/projects-images/rose_manual.pdf
8. KASPRZYCKA K., KOBAKA W.: Właściwości psychofizyczne operatora w ocenie kwalifikacyjnej kandydatów na kontrolera ruchu lotniczego (Psychophysical properties of the operator in assessing the qualification of candidates for the air traffic controller). Job Mgr under the direction of J. Gruszecki, PRz, 2013.