Journal of Theoretical and Applied Information Technology
15th January 2020. Vol.98. No 01

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENTROPY BASED DECISION MAKING METHOD IN MANAGING THE DEVELOPMENT OF A SOCIO-INFORMATIONAL SYSTEM

[1] V. KULIKOV, [1]K. IKLASSOVA, [1]A. KAZANBAYEVA
[1] Department of Information and Communication Technologies
M. Kozybayev North Kazakhstan State University, Petropavlovsk, Kazakhstan
E-mail: [1]qwertyra@mail.ru, [1]kiklasova@mail.ru, [1]akazanbaeva83@mail.ru

ABSTRACT

This article describes the informational and entropic approach to informational flow analysis in management system of higher educational establishment. A review of a range of theoretically suitable entropy versions has been presented as well as the facts with calculations examples, oriented to the further application to decision-making system in management of higher educational establishment workflow processes. This study contains the new method for a system of support of management decision making in prioritizing and choosing of the ways of building university educational policy. Applicability of receiving of a new data that can improve regional educational system by evaluation of structuring of curriculums and educational programs.

Keywords: *Information; Entropy; Decision-Making; Management; Higher Education.*

1. INTRODUCTION

The problems of managing complex systems are associated with an excessive number of parameters characterizing the state of the system. Collecting and processing data on all parameters is a very complex and expensive procedure. However, it is impossible to refuse it because of the need to obtain relevant and reliable data for the adoption and implementation of management decisions. Thus, it is necessary to look for ways and means to reduce the number of monitored parameters, create algorithms to predict the presence of undesirable processes in a controlled system, and develop recommendations for a more detailed analysis of individual systems (or groups of systems) of the environment.

Decision-making process in management of complicated systems in socio-economic sector, particularly – in education sector, on frequent occasions takes place in uncertainty and directly related to analysis of great number of different-type data. As a rule, uncertainty is caused by the following factors: contradictory, incompleteness or the lack of information; data multidimensionality that describes the object of management; latent interrelation of management system elements. In order to eliminate the uncertainty, various expert, statistic or probabilistic methods cannot be applied in multidimensional data analysis and consideration of the source data as a system of characteristics of management object. This task can be solved due to the entropy calculations and models [1].

Entropic approach was considered in studies on decision-making process in various sector of research: this article [2] describes the entropy functionality for merging and checking of decisions in images analysis and application of computer vision; [3] decision-making method is based on cross-entropy, which allows considering far more information for evaluation of investment alternatives; authors [4] study the attributes weight, used for aggregation of preferences of person that makes decision, also the example has shown which shows the house choosing process with the use of entropy method; hybrid method of making decision in choosing of suppliers with the use of weight entropy; authors of this article [6] study the problems of efficient support in decision-making process in psychological health sector which includes such important parts as calculation of system stability and entropy indicators, received from structural changes; dimensional geographical systems [7] uses entropy for data classification; [8] information entropy can be used for calculation of weight of various criteria for ranking of water resource management scenarios. Entropy can be used: for analysis of requests from decisions support system [9]; as stochastic requests optimizer [10]; for description of machine state

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

properties characteristic [11]; for objective determination of weight criteria [12, 15]; for investment portfolio optimization [13]; for quantitative evaluation of methylation stochastic behavior in certain WGBS samples [14].

Analysis of sources of research for higher education sector reflects the presence of the certain interest in this direction, for example: authors [16] for efficient using of resources in higher education sector study the decision-making process algorithms with the use of relational decisions and determination of entropy weight; curriculums content improvement [17], evaluation of lecturers work in universities [18]; evaluation of information literacy [19].

Knowledge management in higher educational establishments and strategies of knowledge management were studied [20], this approach is based on interviewing and thematic analysis, this article contains [21], [22] comparative analysis of curriculums oriented towards knowledge management with the use of text analysis tools.

Conducted analysis shows that now the problem of formation of a certain set of university curriculums ensuring sustainable development of a higher educational establishment is still open.

A functioning of the socio-economic system on the HEE example is analyzed in this article. We are positioning the monitoring of indicative plan of strategic development as a process, which requires calculations of quantitative characteristics of information. These characteristics may be reflected in the evaluation of the entropy of many documents that are necessary for monitoring, and can be defined by amount of data that supports the managerial decision-making [23].

Indicative control implies the monitoring of indicators; comparing forecasting data and goal indicators; evaluation and selection of the best alternatives of development and the most efficient alternate solutions.

Indicators are multidimensional and of different-type, for example [24]:

- statistic indicators, that are representatives of level of internationalization (number of foreign students), including those who studies on a paid basis; number of educational programs in English; international students mobility indicators etc.);
- Qualitative evaluation in national and international ratings (being among the top-10 best comprehensive HEEs in Kazakhstan according to the national rating and others);
- Economic efficiency indicators (energy saving; financial viability indicators; income diversification etc.);
- «procedural» indicators (changing the legal status; corporate management body functioning etc.).

Recently, many publications have been issued on improving the quality of education and the need to introduce unbiased assessment of this quality. Often, assessments of the training/education quality are considered both in a simplified and traditional way [25-31]. The discourse is mainly related to assessing only the individual achievements of students and to the procedures of licensing, certification and accreditation of educational institutions. The problem of assessing the quality of higher education is much deeper and serious. This article proposes to form characteristics of the educational achievements quality through an entropy approach.

The system of testing the training results has been transformed in the direction of quality control of education. It should be emphasized that the introduction of innovative educational technologies in continuing education is becoming an increasingly effective tool for professional development.

Tests can be used at any stage of training. Some of them are designed to assess the readiness of students to learn a new training course, others are aimed to identify specific gaps in knowledge of students in order to plan the necessary targeted corrective work, and some are used to predict the further learning process. Basic testing includes input, current and progress checks, and interim certification for each discipline, including external evaluation of educational achievements (VOUD) and state final certification.

2. MATERIALS AND METHODS

For complex stochastic systems, including socio-economic, entropy is a characteristic of the dynamic stability of the system - a key property of the system from the point of view of the system approach and the general theory of systems. This makes it possible to use entropy as an integral parameter for monitoring the state of the socio-economic system.

The difficulties in applying entropy as an integral parameter in its classical interpretation in relation to real socio-economic systems are mainly due to the impossibility of obtaining its exact value, which would systematically take into account all the factors affecting the dynamics of the system. However, an indirect estimation of entropy is possible by estimating the rate of change of system parameters.

Obviously, when any of the parameters changes, the amount of information necessary to describe the current state of the system changes, therefore, its entropy changes. The rate of change of system

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

parameters, and hence the rate of change of entropy, characterizes the dynamics of the system relative to the external environment. A robust estimate of the rate of change makes it possible to predict the future state of the system using classic technical analysis tools. The reverse is also true: when the entropy changes, the amount of information in the system description changes, therefore, the parameters of the system description have changed. Moreover, if the rate of change of the system parameters does not correspond to the rate of change of the external environment, the system leaves the state of dynamic equilibrium. This is evidenced by a critical change in the output parameters of the system and their deviation from the predicted values. The property of additivity of entropy allows us to identify deviations in individual elements of the system based on the analysis of a phenomenological model.

Indicative plan values behavior are monitored through the workflow of university departments. Documents are incited by i subject of workflow and submitted to another subject j. Each copy or the fact of interaction (endorsing, registration etc.) will be taken into account individually, up to and including the recording of the time of observation of a significant event.

Let us denote the quantity of forwarded documentation in terms of socio-informational system $n_{ij}(t)$, all documents range, available in this system at t-moment - $N(t)$. Determining the intensity

$$p_{ij}(t) = \frac{n_{ij}(t)}{N(t)} \quad (1)$$

now we can form the desired system behavior as the tendency of entropy (S) change of observable workflow in the line of extremal values of entropy $S \rightarrow extr$ under compliance with all laws of system evolution (so-called «frame» that determines behavioral constrains of the system).

As a rule, systems where the workflow determines the behavior are accompanied by indicators monitoring that describes the content side, expressed in quantitative form. It is arguable that there is always a set of indicators in the volume of R, or set of document flow functions,

$$f_r(x_{ij}),\ r=\overline{1,R}, \quad (2)$$

where x_{ij}– some of characteristics of workflow intentionality between subjects i and j of described system. These characteristics x_{ij} presented in consolidated form are of our prime concern, as a rule, we use them in the purpose of making system management decisions. In addition, the average values E_r are known (or assigned) and taken as indicative ones, for example. Then the goal for the monitoring of indicators can be presented in optimization model form:

$$S \rightarrow max, \quad (3a)$$

$$\sum_{ij}^{N} p_{ij} f_r(x_{ij}) = E_r \text{ for each } r, r = \overline{1,R} \quad (3b)$$

$$\sum_{ij}^{N} p_{ij} = 1 \quad (3c)$$

Solution of this model will provide the optimal (in terms of workflow entropy evaluation) plan of workflow intensionality, reflecting the reaching of indicators for the system.

Generic solution of stated optimization model may be found by using Lagrange function with undetermined coefficients α, β_r (all R indicative requirements are taken into account):

$$L(p_{ij}, \alpha, \beta_r) = S - \alpha\left(\sum_{ij}^{N} p_{ij} - 1\right) - \sum_r \beta_r \left(\sum_{ij}^{N} p_{ij} f_r(x_{ij}) - E_r\right) \quad (4)$$

In order to find the optimal distribution p_{ij} we are using the entropy extremum criteria and obtaining the set of differential equations

$$\frac{\partial L}{\partial p_{ij}} = 0 \quad (5)$$

Entropy S can be used in the sense of Shannon, Renyi or Tsallis [24].

For example, in the case of Shannon entropy

$$S = -\sum_{ij}^{N} p_{ij} \log p_{ij} \quad (6)$$

and only one indicator r=1, we get the following for each i and j:

$$\frac{\partial L}{\partial p_{ij}} = -\log p_{ij} - 1 - \beta f(x_{ij}) - \alpha \quad (7)$$

$$-\log p_{ij} - 1 - \beta f(x_{ij}) - \alpha = 0 \quad (8)$$

$$p_{ij} = e^{-1-\alpha-\beta f(x_{ij})} \quad (9)$$

Whence it follows that after summation with provisions for equation (2c), we get:

$$\sum_{ij}^{N} e^{-1-\alpha-\beta f(x_{ij})} = e^{-1-\alpha} \sum_{ij}^{N} e^{-\beta f(x_{ij})} = 1 \quad (10)$$

or

$$e^{\alpha} = \sum_{ij}^{N} e^{-\beta f(x_{ij})-1} \quad (11)$$

The complexity of Lagrange coefficients interrelation research is determined by type *f(·)* non-trivial dependence, so that the only indicator *E* has been introduced. Having got a pair of Lagrange coefficients calculated, for example, by using approximative methods, it is possible to uniquely establish all workflows in a stably functioning workflow system at every instant, including the forecast how planned indicator *E* will be reflected on each performer`s workload.

In case of using Renyi entropy

$$S_R = \frac{1}{1-q} log \sum_{ij}^{N} p_{ij}^{q} \quad (12)$$

if r=1, then:

$$\frac{\partial L_R}{\partial p_{ij}} = \frac{1}{1-q}\frac{p_{ij}^{q-1}}{\sum_{ij} p_{ij}^{q}} - \alpha - \beta f(x_{ij}) = 0 \quad (13)$$

Multiplying by pij, summed iand j, we obtain:

$$\sum_{ij}^{N}\left(\frac{1}{1-q}\frac{p_{ij}p_{ij}^{q-1}}{\sum_{ij} p_{ij}^{q}} - \alpha p_{ij} - \beta p_{ij} f(x_{ij})\right) = 0 \quad (14)$$

$$\frac{1}{1-q} - \alpha \sum_{ij}^{N} p_{ij} - \beta \sum_{ij}^{N} p_{ij}\, f(x_{ij}) = 0 \quad (15)$$

As (2c), the, denoted

$$E = \sum_{ij}^{N} p_{ij}\, f(x_{ij}) \quad (16)$$

we shall finally obtain:

$$\frac{1}{1-q} = \alpha + \beta E \quad (17)$$

Indicators interpretation:
E – indicator;
q – Renyi parameter;
α –Lagrange coefficient for taking intensionality into account;
β –Lagrange coefficient for taking indicator into account;

Determined formula for (*q, E,* α, β) interrelation literally means the following: the only workflow indicator can be replaced by Renyi entropy parameter.

Besides, if *q*=0, entropy gets the «structural» form and virtually abbreviated with the formula [32]

$$S=\log N \quad (18)$$

This could be Hartley equation, if we used equiprobable distribution p_{ij}. However, here it is only a coincidence caused by the "successful" partitioning of the structure of the workflow system.

Let us suppose that if *q*=0 for this system we can introduce such partitioning scale parameter ε, so we can associate it with indexes, pointing to the performers of documents of different scale (on behalf of the organization, department, service, etc.). The content "number of documents produced outside a subdivision of any scale corresponds to the representation of each performer in this subdivision" formally means the existence of a numerical limit indicating the generalized fractal Renyi dimension:

$$d_0 = \lim_{\varepsilon \to 0} \frac{S}{ln\left(\frac{1}{\varepsilon}\right)} = \lim_{\varepsilon \to 0} \frac{ln\left(\sum_{i=1}^{N(\varepsilon)} p_{ij}^{0}\right)}{ln\left(\frac{1}{\varepsilon}\right)} \quad (19)$$

In such a way we shall get Hausdorff dimension, which determines fractality of workflow system structure [25].

In case of using the formula of Tsallis entropy

$$S_T = \frac{1}{q-1}\left(1 - \sum_{ij}^{N} p_{ij}^{q}\right) \quad (20)$$

instead of previously used Shannon, without resource to the only indicator premise, then we get:

$$\frac{\partial L_R}{\partial p_{ij}} = \frac{q}{1-q} p_{ij}^{q-1} - \alpha - \sum_r \beta_r f_r(x_{ij}) = 0 \quad (21)$$

We get the parametric equation (use interpretation in case of Shannon entropy)

$$\sum_{ij}^{N}\left(\frac{1-q}{q}\left(\alpha + \sum_r \beta_r f_r(x_{ij})\right)\right)^{\frac{1}{q-1}} = 1 \quad (22).$$

3. RESULTS AND DISCUSSION

3.1 Entropy Assessment As A Decision-Making Method In Document Management

Let us consider as a calculated example the supporting documents of M.Kozybaev North Kazakhstan state university curriculum.

Specialty curriculum is considered to be the curriculum base document – 56. However, the constant higher-education system reformation lead to the annual curriculum changing, so theoretically the number of documents increased to 562 (4 courses of intramural, extramural, accelerated academic programmes), developed by 26 administering sub-departments for 56 specialties.

Let us take into account that the number of plans can be "reduced" to the number of study groups 360+201=561, as because of the birth dearth the size modern batch virtually equals to the size of study

Journal of Theoretical and Applied Information Technology
15th January 2020. Vol.98. No 01

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

group. In fact: 8 faculties form $\sum n_{ij}$=339=N Bachelor`s curriculums (26 sub-departments, 56 specialties, some of the curriculums are the «translations» to the corresponding language of study) shown in table 1.

Table1: University data summary

Departments	Number
Faculties	8
Sub-departments	31
Bachelors specialties	56
Intramural study groups	360
Extramural Study Groups	201

It should be noted that the number of specialties (curriculums) is changing every year, even though the quality of documents should remain suitable both to HEE opportunities, and state standards. Thus, the requirements of the balance of the credits volume and the teaching materials content on subjects, on the basis of which $f(x_{ij})$ are generated. Here is the calculated value of classical information Shannon entropy:

$$S = -\sum_{ij}^{56} p_{ij} \, log \, p_{ij}, \sim 5.64312680,$$

by taking $p_{ij}=n_{ij}/N$.

In case of low q Renyi entropy is taken as structural, so we shall evaluate it if q=1:

$$S_R = \frac{1}{1-q} log \sum_{ij}^{56} p_{ij}^q, \sim 5.64312680.$$

If q=1 Renyi entropy equals to Shannon entropy, however, if q-value increases – the value of Renyi entropy declines.

Tsallis entropy be haves like q-value:

$$S_T = \frac{1}{q-1}(1 - \sum_{ij}^{56} p_{ij}^q), \sim 3.9115 \text{ by taking } q \sim 1.$$

As known, Tsallis entropy, reaches its maximum (for q>1) only in case of equally probable distribution, and equals to

$$S = \frac{N^{1-q}-1}{1-q} \quad (23)$$

if $p_{ij}=1/N$ (for the example above it is 1/339, N=339) it will be equal to Shannon entropy S~5.64312680 by taking q~1.

Tsallis entropy offers the greatest promise for complicatedly interacting workflow (but non-additive as observed with Shannon and Renyi entropy), as it takes into account the correlation (interaction) of elements with subsystems. To take into account this complicated interaction, curriculum should be divided into modules and blocks, remembering that the compulsory subjects of State Compulsory Educational Standard are offered for batches, and elective subjects, following "fashionable educational" trends, are changing over the years from plan to plan, etc. It illustrates horizontal and vertical interrelation, supposing integration into learning batches, if credit volumes are pertinent and periods coincide. However, the calculation noticeably becomes more complicated, as the variability of elective subjects significantly diversifies the curriculums, even those combined in specializations.

Let us consider the curriculum of the university for the current year:

S1 – 339 current curriculums are presented in the new HEE structure (reorganization of faculties and some sub-departments),

S2 – 338 curriculums are presented before faculties reorganization, presented as "old" structure of the workflow;

S3 – 328 curriculums are presented, curriculums for the first open specialty courses are not included (new structure of faculties);

S4 – 280 curriculums are presented in new structure of faculties; curriculums of graduation courses that corresponded to the workflow of the "old" HEE structure are not included

S5 – 56 basic curriculums are presented, correspond to the number of university specialties; with any workflow structure, are theoretically minimally possible and ideal ("symmetrical"), but unrealistic (virtual).

Values of structural Renyi entropy are presented in table 2.

Table 2. Renyi entropy-q-value interrelation

q	S1	S2	S3	S4	S5
0	5,807	5,615	5,644	5,755	5,807
1	5,643	5,496	5,536	5,610	5,807
2	5,545	5,408	5,464	5,509	5,807
3	5,480	5,340	5,411	5,434	5,807
4	5,432	5,283	5,369	5,376	5,807
5	5,394	5,236	5,335	5,328	5,807

Renyi entropy-q-value interrelation diagram is presented in figure 1.

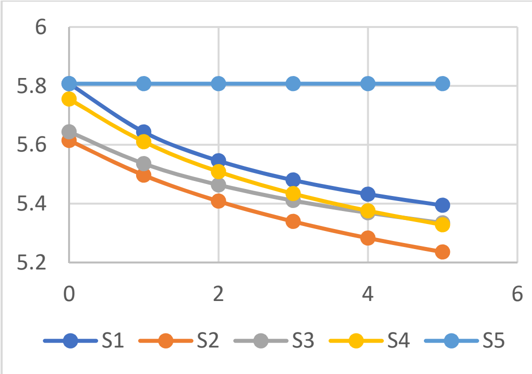

Fig. 1 Renyi entropy-q-value interrelation diagram

All range of change of entropy valuations fall with the limits between “ideal” maximum value of variant *S5* (observed in the “symmetric” case and equals to the Shannon valuation) and minimum value, corresponding to case *S2* (“old” structure of faculties and accordingly workflow).

On the example of the case *S3* and *S4*, the possibility of inversion of the Renyi entropy valuation with q-value increasing. In the case of this pair, in the interval of changing the value of parameter q from 4 to 5, the evaluation of the contribution of the previous curriculums changes because of graduations and lack of enrollment in specialties with the “old” curriculums (*S3*), as opposed to contribution of the “new” curriculums which do not provide the enrollment in past years (*S4*).

This circumstance is extremely important, especially in connection with the aforementioned possibility of interpreting the only figure (indicator) of workflow as a parameter of Renyi entropy. Figure that was introduced for workflow quality characteristic on a *q*-level of “information diversity” will be evaluated in another way on a *q*-level of conditional “correlation”. It should be noted that: comparison analogy is not to be irrelevant in passing from average to medial indicators.

The observed increase of “diversity” evaluation level of system *S1* in comparison with *S2* for the range of Renyi indicators convinces of:

- some “information advantage” of the new structure of faculties (general HEE workflow contains virtually the same number of documents plans);
- need for completely different indicators characterizing the workflow.

The given calculation example illustrates the following: the volume of workflow cannot reflect the structural diversity, in contrast to the entropy evaluation.

For example, the following interpretation of *q*-parameter in Renyi entropy is acceptable having the only indicator E.

Previously derive formula (17) interrelate *E* and the increasing of

$$\frac{1}{1-q} \qquad (24)$$

Whence it follows that *q*-value shall be less than unity, but shall tend to be closer to unity as indicator grows, which let us to replace the Renyi entropy with the Shannon one. And in case of Shannon entropy we can refer to the common used interpretation whereby we shall use quantitative evaluations according to the workflow volume, what is more these volumes (*i, j, N*) will be indirectly taken into account, using the Lagrange coefficients, considering the contribution of workflow structural constrains. It can be assumed how significant the structural constraints in this case can be. For example, the presence of a single holder of a ”round seal” affixed to each copy of the curriculum after “reconciliation” assumes that the limitation will be the working time limit of this “seal holder” allowed for the approval of such documents within the framework of labor legislation.

For the same data, the calculation of the Tsallis entropy showed the following results.

Table 3. Tsallis entropy-q-value interrelation

q	S1	S2	S3	S4	S5
0,3	21,8	20,2	19,6	20,1	19,9
0,6	9,6	9,2	8,6	8,93	9,1
0,9	4,7	4,6	3,62	3,7	4,6
1,5	1,7	1,7	1,7	1,7	1,6

Tsallis entropy-q-value interrelation diagram is presented in figure 2

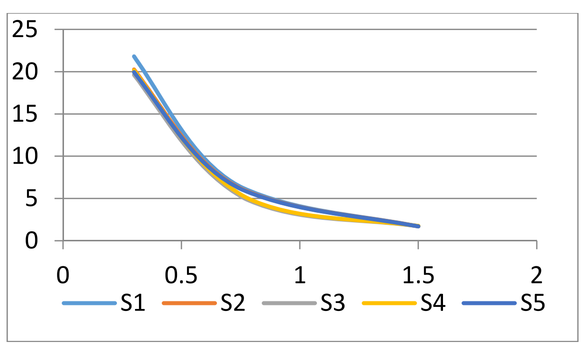

Fig. 2 Tsallis entropy-q-value interrelation diagram

On the basis of data obtained, conclusions can be made regarding the level of system organization, the complication and quality of the management tasks solved in decision-making process. The development of a system changes its entropy, progresses/degrades uncertainty and orderliness. This phenomenon can be interpreted as a process of changing opportunities. These processes can become the basis for the well-timed reorganization of the workflow structure, and for the opening up of new possible trajectories for further system development.

3.2 Formation Of The Quality Characteristics Of Educational Achievements Using The Entropy Approach

For any educational institution of the Republic of Kazakhstan (not statistically confirmed but derived from the personal experience of the authors), model assumptions are specified as follows:

A) As a rule, all three control subsystems: input (dropout), current (regular) and output (final) are implemented. Here we should note that "unloading" of the process recording for data processing, as a rule, represents a separate problem.

B) As a rule, "elimination of ballast" documents are either not performed or does not have any cause-effect relations with ensuring the declared quality of subsystems functioning.

C) As a rule, the pool of documents passively reaches some level of redundancy in anit-happened-accidently way.

Let's consider the subsystem of External Evaluation of Educational Achievements (VOUD):

- 4 sets of tasks for 25 questions;
- each question is assigned to groups of 5 to 8 potential answers;
- each group contains 1 to 3 answers marked as "correct";
- the result of the VOUD for the certified persons shall be the accrual of points for the selection of marked options in the range from 0 to 200.
- the only criterion of (16) is the condition: the proportion of Z subjects must score at least W points, which can be interpreted as the "average" set of points for the certified points not lower than the threshold of E≥W.

We shall assume that the values of Z and W do not change with the increase in the volume of subjects (it is consistent with the VOUD practice), and the function $f(x_{ij})$ is as follows:

- argument x_{ij} plays the role of the "meeting" event of the certified lot with a sample of tests;
- substantially $f(x_{ij})$ represents the result for the i^{th} lot of the certified persons from all test pool N (the lot can consist of the only certified person) of the j^{th} sample of the test documents made by rules of an external assessment of educational achievements;
- function x_{ij} from each x_{ij} is non-negative.

Then $f(x_{ij})$ can be assumed constant, that is equal to E (E≥W), which follows from (16) taking into account the normalizing constraint (3c).

For example, such interpretations of "almost incredible" VOUD certification processes are possible:

I) the situation is formed in an unnatural way, through the artificial creation of the certified lot. In particular, all unable (unwilling) to be certified "receive" sick leave instead of the VOUD test. We have a situation of equal $f(x_{ij})$ (if those capable to pass VOUD test answer only to a threshold level E). This corresponds to a situation of equal p_{ij} for Shannon-Hartley entropy function (18) (interpreted as extremely undefined), but it leads to the guaranteed E≥W. In other words, "equally possible" testology approach gives, however, a guaranteed result.

P) "dummy" VOUD certification. There is only one p_{ij} value other than zero (and equal to one): the process has an absolutely regular basis, since even the entire test pool presented to a single group with x_{ij} characteristics of "test experts" (assisting the certified persons in an unauthorized way) always gives an indicator equal to the expert E≥W.

However, the elimination of I) and P) situations (i.e. the ability to "clean up lists of certified persons with sick leaves" and/or actually be certified be the "experts") will lead to a change in the values of $f(x_{ij})$ and in the evaluation of entropy. Indeed, any distribution of p_{ij}, different from uniform, provokes non-constant $f(\cdot)$, since nonzero E gives correlation of$p$and$f(\cdot)$ in the interpretation of$p$and $f(\cdot)$ as random variables. The substantive requirements of VOUD include the requirement of hitting two-dimensional random variable

$$E(\cdot,\cdot) = \sum_{ij}^{N}\left(p_{ij} f(x_{ij}) \stackrel{\text{def}}{=} E_{ij}\right) \qquad (25)$$

in the ellipse of dispersion [16], parameterized by Z and W values.

The analyzed above I) and P) cases of VOUD certification processes in the proposed interpretation perform an artificial reduction in the variance of the scatter of results with the mean shift in the direction of the maximum. Therefore, I) and P) can be recognized by comparing the entropy values calculated from both the session results and the VOUD results. The "artificiality" of the I) situation is equivalent to the "master mindness" of the P) situation (and both are different from the session

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

situation when, *all* are attested *independently* one way or another), while the role of entropy in the first shall belong to Renyi entropy with q-parameter calculated from the level of E index, based on W

$$E = \sum_{ij}^{N} p_{ij}\, f(x_{ij}),\ (E \geq W) \qquad (26)$$

Let's consider a model example.

A small group of 3 people takes a session:

- every appraisee answers to 100 questions out of 300 questions in the total base volume;
- at the same time, 34, 33 and 33 questions have 5-8 answers with, respectively, one, two and three "correctly"- marked answers;
- the documents of the test record are atomic (the atom is equal to one question), and:

$\sqrt{p_{ij}} \in \{0, 1/300, 2/300, 3/300\}$,

a $f(x_{ij}) \in \{0, 1/3, 2/3, 1, 4/3, 5/3, 2\}$ for a session;

$\sqrt{p_{ij}} \in \{0, 1/200, 2/200\}$, a $f(x_{ij}) \in \{0, 1/2, 1, 3/2, 2\}$ for VOUD (C-student is "hidden");

- a "C-student", who always puts only one check-mark, barely gains a total of 100 points and passes the session, is not involved in VOUD (has a supporting document);
- a "B-student" never selects three marks, passes the session and VOUD;
- an "A-student" is not afraid to lose points by choosing "extras" (wrong answer), hoping to score the highest possible score, successfully passes the session and VOUD.

Summarizing $\sum_{ij}^{300} p_{ij} f(x_{ij})$, we can calculate:

SE) Shannon Entropy $S = -\sum_{ij}^{300} p_{ij} \log p_{ij}$.

The value of the Shannon entropy:

- for the session (calculations for three participants) is in the range from log(300)≈8.23 (limit case when everyone answered different questions) to 3log(33,(3))≈15.18 (limit case when all answered one set of questions);
- for VOUD (similar calculations for two participants) from log(200)≈7.65 to 2log(50)≈11.29.

So, intervals [8.23; 15.18] and [7.65; 11.29] give overlapping, so we have:

- the session situation indistinguishable from that of VOUD's, if p_{ij} distribution s are completely independent;

If p_{ij} s are of the same type, for example, the questions did not match with anybody, the "session" will give the value of the entropy strictly greater than the entropy of VOUD.

RE) Renyi Entropy $S_R = \frac{1}{1-q} \log \sum_{ij}^{300} p_{ij}^{q}$.

Similar calculations give intervals of [3.47; 8.23] and [4.64;7.65], respectively, for the session and for VOUD, and an indistinguishable situation, except for "the same type" (as for Shannon entropy).

We shall note that the upper Renyi-evaluation of intervals does not depend on q and coincides with the lower Shannon-evaluation, which is obtained for the minimum possible set of documents (test items) – when all test takers receive the same 100 questions (such evaluation estimates give the results of a trial VOUD-testing on a single variant of tasks).

TE) Tsallis Entropy $S_T = \frac{1}{q-1}(1 - \sum_{ij}^{N} p_{ij}^{q})$.

Tsallis entropy, as is known [26], reaches a maximum of $S = \frac{N^{1-q}-1}{1-q}$ for q>1 in the case of equally probable distributions, that is at $p_{ij}=1/N$ for the model example. Value of the entropy:

- is N=300,S≈8.23 for the session at q≈1;
- is N=200, S≈7.65 for VOUD at q≈1.

If the questions in the test at least partially coincided, Tsallis entropy drops markedly up to S≈4,61both for the session and for VOUD.

In this case the situation looks similar to the one described above, as variety of combinations of tasks reduces at VOUD (part of the test books will not be distributed for the reasons mentioned above). It may be worth splitting the test task into a documentary set of question-answer pairs, taking into account the fact that some pairs will require a meaningful "hint" because of the internal dependence of the answer options previously combined with the question. Tsallis entropy with q≠1parameter is usually applied specifically for such options.

The "entropy" scheme of replacement of the VOUD concept (or other external check of residual knowledge) by results of the session results analysis is proposed here.

The value S_R^* of the Renyi entropy is chosen from the internal points of the overlapping intervals, for example, on the basis of correspondence to the value of E indicator. If talking about the model example, this shall be not lower than 4.64, and the possible choice of the upper level equal to 7.65 allows the variation of q parameter within a wide range without fundamental influence on the result.

A pool of documents is collected from the database of test tasks used at the session, the statistics on which gives exactly the selected entropy value S_R^*.

The results of the appraises are surveyed on the constructed pool.

The session databases contain the entire range of the appraises (there is no effect of replacing the test with a sick leave), and there is no specially selected group of "experts" that "enhances" the results (the testing during the session is under video surveillance; moreover, it is possible to effectively record the process by side sources of information).

The results on the sample pool are averaged according to the VOUD requirements.

Repeated testing in the form of "VOUD-session" on the session basis at the graduation course and recording among others new (repeated) results according to the proposed scheme (1.-4.) shall assess the sustainability of the results following the criterion of the "residual" knowledge.

4. CONCLUSION

In this article the use of entropy as a system structure index as well as the level of interrelation of internal information processes through the example of socio-economic system – higher educational establishment has been found.

In examined case, Renyi entropy is being considered the significance test, where by the optimal non-uniform distribution is established, and provides the equilibrium of set of interacting elements of the system.

Information support in management decision-making process aimed at the system development can be based on statistic information concerning the workflow between system elements. Even more, the decision to be supported by data, characterizing the ratio between orderliness, structure index and controllability of the system.

As indicative plan indicators monitoring can be achieved by analyzing the workflow of HEE structural departments, the workflow should be characterized not only by the number of documents (such indicators cannot identify the difference where it is), but be supplemented by an entropy valuation as an indicator reflecting the change of structural possibilities in management of students` curriculums.

Further research is oriented towards studying of opportunity of implementation of this approach in a working information system of university for expanding of intellectual methods of making decisions. Result presented in this article can become the key to successful management of such a complicated system as university.

We shall note here (not statistically proven) that since de facto the VOUD database consists of session materials of the main universities of the Republic of Kazakhstan used for their certified enrollment, the scenario (1.-4.) actually occurs (in the absence of entropy evaluations there is no guarantee that the scheme of the most unsuccessful version of the sample/test pool is not implemented).

We shall emphasize that the model calculation is given only for the ease of checking the calculations, all the conclusions and suggestions do not change for any more voluminous and realistic situation.

Indicators reflecting the content of the control in an average quantitative form are formed in systems of knowledge control. They are usually given as functions of the characteristics of the processes and explicitly or implicitly take into account or tend to take into account the "typical" manifestations, assuming data with stable characteristics that are statistically significant and sufficient for the analysis. The paper shows the availability and possibilities of other methods of analysis, based primarily on entropy characteristics, less demanding on the volume and quality of data and, at the same time, actively using, for example, materials already available in the system and the generally accepted document flow. New entropic approaches may be used in the control system for decision-making on its improvement.

REFRENCES:

[1] I.V. Prangishvili, "Entropic and other system regularity", *Complicated systems control issues, Ed. Moscow, Russia: Nauka*, 2003.

[2] O. Gunay, B.U. Toreyin, K. Kose, and A.E. Cetin, "Entropy-Functional-Based Online Adaptive Decision Fusion Framework With Application to Wildfire Detection in Video," *IEEE Transactions On Image Processing*, 21(5), 2012, 2853-2865.

[3] J. Ye, "Cross-Entropy of Dual Hesitant Fuzzy Sets for Multiple Attribute Decision-Making", *International Journal Of Decision Support System Technology,* 8, 2018, 20-30.

[4] H. Garg, N. Agarwal, A. Choubey, "Entropy Based Multi-criteria Decision Making Method under Fuzzy Environment and Unknown Attribute Weights", *Global J Technol Optim*, 6(3), 2015, 20-30. doi:10.4172/2229-8711.1000182

[5] P.D. Liu, X.Y. Qin, "An Extended VIKOR Method for Decision Making Problem with Interval-Valued Linguistic Intuitionistic Fuzzy Numbers Based on Entropy", *Informatica*, 28, 2017, 665-685.

[6] C.R. García-Alonso, N. Almeda, J.A. Salinas-Pérez, M.R. Gutiérrez-Colosía, J.J. Uriarte-Uriarte, L. Salvador-Carulla, "A decision support system for a ssessing management interventions in a mental health ecosystem: The case of Bizkaia (Basque Country, Spain)", *PLoS ONE*, 14(2), 2019.

[7] S. Wan, T. C. Lei, and T. Y. Chou, "Optimized object-based image classification: Development

of landslide knowledge decision support system," *Arabian Journal of Geosciences*, 7(5), 2014, 2059-2070.

[8] M. Karamouz, A. Ahmadi, M. S. S. Yazdi and Behzad Ahmadi, "Economic assessment of water resources management strategies," *Journal of Irrigation Drainage Engineering*, 140 (1), 2014.

[9] M. Sharma, G. Singh, R. Singh and G. Singh, "Analysis of DSS queries using entropy based restricted genetic algorithm," *Applied Mathematics and Information Sciences*, 9(5), 2015, 2599-2609.

[10] M. Sharma, G. Singh and R. Singh, "Design and analysis of stochastic DSS query optimizers in a distributed database system," *Egyptian Informatics Journal*, 17(2), 2016, 161-173.

[11] Z. Su, H. Xiao, Y. Zhang, B. Tang and Y. Jiang, "Machinery running state identification based on discriminant semi-supervised local tangent space alignment for feature fusion and extraction," *Measurement Science and Technology*, 28(5), 2017.

[12] L. Zhong and L. Yao, "An ELECTRE I-based multi-criteria group decision making method with interval type-2 fuzzy numbers and its application to supplier selection," *Applied Soft Computing Journal*, 57(8), 2017, 556-576.

[13] H. Jalota, M. Thakur and G. Mittal, "A credibilistic decision support system for portfolio optimization," *Applied Soft Computing Journal*, 59(10), 2017, 512-528.

[14] G. Jenkinson, J. Abante, A. P. Feinberg and J. Goutsias, "An information-theoretic approach to the modeling and analysis of whole-genome bisulfite sequencing data," *BMC Bioinformatics,* 19 (3), 2018.

[15] L. Zhang, and H. Zhou, "A group decision-making model based on vague sets for emergency improvisation decision," *International Journal of Applied Decision Sciences*, 9(3), 2016, 283-306.

[16] J. Huan and R. Bo, "Application of the Big Data Grey Relational Decision-Making Algorithm to the Evaluation of Resource Utilization in Higher Education," *International Journal Of Enterprise Information Systems*, 14(2), 2018, 43-55.

[17] A. A. Toptsis, "Toward a better syllabus: Entropy-driven introspection for alternative lesson plans," *Journal of Computing and Information Technology*, 17(4), 2009, 335-346.

[18] N. Hein, A. Kroenke and M.M.R. Júnior, "Professor assessment using multi-criteria decision analysis," *Procedia Computer Science*, 55, 2015, 539-548.

[19] Y. Wang, H. Li and Z. Ding, "Information Literacy Assessment with a Modified Hybrid Differential Evolution with Model-Based Reinitialization", *Computational Intelligence and Neuroscience*, 22, 2018, 1-21.

[20] P. V. Ramjeawon, J. Rowley, "Knowledge management in higher education institutions in Mauritius," *International Journal of Educational Management*, 32(7), 2018, 1319-1332.

[21] P. V. Ramjeawon, J. Rowley, "Knowledge management in higher education institutions in Mauritius", *International Journal of Educational Management*, 32(7), 2018, 1319-1332. doi: 10.1108/IJEM-05-2017-0129.

[22] H. Frank Cervone, "What does the evolution of curriculum in knowledge management programs tell us about the future of the field?", *VINE Journal of Information and Knowledge Management Systems*, 47(4), 2017, 454-466. doi:10.1108/VJIKMS-12-2016-0072.

[23] V. Sochnev, "Aspects of using the informational approach for company management", *International symposium "Sovnet-99" studies*, 1999.

[24] *V. Kulikova, K. Iklassova, A. Kazanbayeva,* "Development Of A Decisionmaking Method To Form The Indicators For A University Development Plan", *Eastern-European Journal of Enterprise Technologies*, 3(99), 12-21, 2019. doi: 10.15587/1729-4061.2019.169193.

[25] A.G. Wilson, "The Use of the Concept of Entropy in System Modelling", *Operational Research Quarterly (1970-1977),* 21, 1970, 247-265.

[26] J. Julliand, O. Kouchnarenko, P. A. Masson, G. Voiron, "Test Generation from Event System Abstraction to Cover Their States and Transition" *Programming and computer software,* 44, 2018, 1-14.

[27] H. Bride, J. Julliand, and P.A. Masson, "Tri-modal under-approximation for test generation", *Sci. Comput. Program.*, 132, 2016, 190-208.

[28] P. Godefroid, "Compositional dynamic test generation", *ACM SIGPLAN Notices*, 42 (1), 2007, 47-54.

[29] P.-C. Bué, J. Julliand, and P.A. Masson, "Association of under-approximation

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

techniques for generating tests from models" *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* 6706 LNCS, 2011, 51-68.

[30] P. Godefroid, N. Klarlund, and K. Sen, "DART: Directed Automated Random Testing" *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI),* (2005) 213-223.

[31] N. Tillmann, J. De Halleux, "Pex-white box test generation for.NET" *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4966 LNCS, 2008, 134-153.

[32] R. Majumdar, K. Sen, "Hybrid Concolic Testing", *ICSE 2007: 29th International conference on software engineering, proceedings,* 2007, 416-426.

[33] O. V. Chumak, "Entropies and fractals in data analysis", *Publication of Sternberg Astronomical Institute Moscow University: Russia, Moscow*, 2010.