Entropy approach to the evaluation of the integration processes in agro-industrial complex

A S Dulesov¹, D J Karandeev¹, T G Krasnova¹, R I Bazhenov² and V P Kochetkov³

¹ Katanov Khakass State University, 92, Lenina ave., Abakan, 655017, Russia
² Sholom-Aleichem Priamursky State University, 70A, Shirokaya street, Birobidzhan, 679015, Russia
³ Khakas Technical Institute of Siberian Federal University, 15, Komarova ave., Abakan, 655017, Russia

E-mail: den_dr_house_1991@mail.ru

Abstract. The possibilities of assessing the integration processes in agro-industrial complex are considered. Preference is given to the study of the influence of network research structures for obtaining, processing and transformation of information. The necessity of development of system recommendations taking into account theoretical achievements and analytical studies based on network analysis and the concept of entropy is indicated. The main directions in the study of integration processes based on the construction and analysis of network structures are briefly considered. The main attention is paid to the entropy approach of evaluation of integration processes. Interpretations of the measure of information uncertainty (entropy) are proposed in order to perform network analysis. Among the preferred models to determine the amount of entropy, the following are proposed: K. Shannon’s measure of uncertainty; cross-entropy and Kullback-Leibler divergence.

1. Introduction

Considering the integration processes in agro-industrial complex (agribusiness), we are talking about vertical and horizontal integration. At the same time, attention is paid to the creation of mechanisms applicable to the introduction of innovation. In essence, it is a question of development of system recommendations with involvement of theoretical achievements and data of the analytical studies based on the network analysis and the concept of entropy. The creation of integrated structures in the conditions of agricultural production has become one of the effective ways to build supply chains, improve the market stability of enterprises and strengthen the competitive position of agricultural producers in the domestic market [1]. Thus, the intensity of the integration process in the agro-industrial complex depends largely on the availability of innovative resources, among which information should be highlighted.

The objectives of integration are fixed in the form of systems of planned activities to develop the potential and capabilities of agricultural enterprises. Considering at the same time modern researches of integration processes, much attention is paid to studying of influence of network structures on development of information, expansion of knowledge and transformation of the latter into new developments and technological processes [2].
2. The main directions of network analysis

The main objects of study in the analysis of integration processes, according to [3], include:

- geographical distribution of the network and the impact of the territorial factor on its integration [4];
- structural characteristics of the network [5] and its management [6];
- cognitive processes between network participants [7] and the strength of connections between them [8];
- analysis of big data and entropy content of processes in networks [9, 10, 11].

Many different areas of research indicate that there is a search for ways and solutions aimed at developing tools that provide assessment and analysis of integration processes. A common feature of all tools is the idea of the importance of the interaction of different subjects.

The evaluation of integration processes should be based on the generation and application of knowledge by the subjects, the success or failure of which depends mainly on the degree of their network interaction and structuredness. Networks provide an opportunity to aggregate heterogeneous and dispersed specific knowledge [12] and, thereby, increase efficiency of its functioning [13].

The growing importance of network analysis for the development of stages of implementation of integration policy due to the expansion of opportunities for the development and application of information technology is justified by the needs of the entities involved in this process. The attractiveness of these tools is related to the strengthening of connection between existing and emerging innovation-oriented networks. Thus, the network analysis is the first stage of work, the results of which will allow decision makers to study the structure, its possible probabilistic state and the degree of uncertainty in obtaining the necessary information.

Uncertainty of states of the network structure is directly related to the presence of factors of the probabilistic nature of behavior [14, 15]. A suitable tool for situations where we don’t know the whole set of causal mechanisms that determine the outcome of a situation can be considered probability theory. Its application is advisable in cases when it is necessary to abandon the idea of predicting specific outcomes and move from a set of deterministic models to probabilistic ones [16]. Their application implies the existence of connections with mathematical statistics. It allows to carry out the first step towards understanding the internal mechanisms of social and economic development and to find connections between individual factors [17] influencing integration processes.

The current statistical database contains retrospective information. The possibility of its use for forecasting processes remains uncertain, so it is often necessary to involve expert assessments that are subjective and in fact are probabilistic in nature. In contrast to the methods of mathematical statistics, mathematical modeling allows you to set the parameters of the connections in the network under consideration between the factors of investment and economic nature [18]. The mathematical model must contain its own parameters, provided that they are constant in time.

The structure of the network cannot be considered to be fully defined, since the nature of the functioning of its internal mechanisms and external factors are endowed with dynamics. Given these circumstances, mathematical statistics can be considered popular, as it suggests the presence of a probabilistic model. It allows you to build a discrete time series, which is a sample of a certain set of data having the properties of stationarity and normality. Therefore, it is possible to estimate the future states of the network that are optimal in terms of accepted criteria.

3. The measure of information uncertainty in the problem of evaluation of integration processes

Considering the probabilistic models and mathematical model of the network state, let's present a brief overview of the possibilities of the measure of information uncertainty (entropy) in the evaluation of integration processes. The presence of the network structure, its retrospective data and model experiments make it possible to determine the entropy, the value of which is necessary to measure the number of potentially achievable states of network elements [19, 20].
The simplest situation with the determination of entropy is possible on the basis of the logarithmic measure of R. Hartley’s information [21]. For example, according to [22] for two \( N \) equal opposite states of the network element, entropy is determined by the expression:

\[ H = \ln N \quad (1) \]

If \( H = 0 \), there is only one state \( N=1 \), which implies its complete predictability, that is, the impossibility of the appearance of another state. For a set of states \( H > 0 \) means that there are several possible states and a lower level of predictability. From the point of view of integration processes, this situation can be interpreted as a range of possible states of the network.

Entropy depends on the probability distribution of the states of the network elements. If the distribution is uniform, then the entropy in magnitude will be maximal. This means that all elements of the network are assigned the same opportunities to participate in the integration process. On the other hand, when the integration capabilities are given to one element of the network, the entropy will be close to zero and, therefore, there is certainty of information about the achievement of the result. The greatest entropy (uniform distribution of states) is equivalent to either extreme uncertainty in obtaining the result from any of the elements, or the maximum probability of waiting for the integration activity from each of the elements of the network. The smallest entropy means either the maximum certainty in obtaining a result from one of the elements (the presence of only one state), or the certainty of information about the achievement of the result.

Let’s to consider mathematical models of entropy determination. In this case, we will use the concept of possibilities as the realization of some random variable \( x \), which takes one of the \( N \) values with probability \( p(x) \). Then the entropy, called "private" will take the form:

\[ H_s = - \ln p(x) \quad (2) \]

The presence of the probability \( p(x) \) in (2) means that the entropy is also a random variable. For each of the states it takes values: \(- \ln p(x_1), - \ln p(x_2), \ldots, - \ln p(x_N)\) with probabilities: \( p(x_1), p(x_2), \ldots, p(x_N) \). If you want to determine the state of the entire network, you can determine the average entropy:

\[ H = \sum_x p(x) \ln p(x) \quad (3) \]

where \( \sum_x p(x) = 1 \)

Expression (3) is the entropy determined by L. Boltzmann in 1877 for thermodynamic system [23] and revived by C. Shannon in 1948 as a measure of uncertainty of information flow [24, 25]. The significance of this expression is essential in the analysis of equal opportunities or the implementation of a random variable with equal probabilities.

One of the important moments in expanding the possibilities of using the entropy paradigm in network analysis was the spread of the notion of entropy for non-probable realizations of a random variable [26, 27]. Determination of the corresponding entropy and the model of its description are associated with the names of the Kullback and Leibler [28]. Let’s consider in some detail the role of this entropy.

Considering the state of the network structure, the value of \( x \) relating to two random distributions is subject to analysis. For example, the first distribution – the planned values of indicators of realization of integration processes in agro-industrial complex, the second – the actual values of the results obtained on the basis of results of a successful integration project on a similar network. Comparing both distributions with each other, it is impossible to say unambiguously whether the probability distributions of \( q(x) \) events coincide with the probabilities of \( p(x) \). The probability distribution \( q(x) \) of the planned indicators is subject to verification and serves as an approximation of the distribution \( p(x) \). The result of the approximation reflects the magnitude of loss (unaccounted for amount) of information in the transition from the desired (expected) distribution \( p(x) \) to the planned \( q(x) \) (which
is due to the ability to manage the integration process. The measure of information considered in this case is called cross-entropy. For two distributions, cross-entropy is determined as follows:

\[ H_p(q) = -\sum_{i=1}^{n} p_i(x) \log_2 q_i(x) \text{, bit} \]  \hspace{1cm} (4)

If we consider equally probable realizations as the occurrence of a single event from the whole set of \( k \) events in the system, the probability of occurrence of a single event is \( p = 1/k \). Since all probabilities are equal, it is possible to speak about the equality of all \( x \) values of the expected distribution \( p_i(x) \). Distributions of this kind are characteristic of a stationary flow of events. Then the cross-entropy according to (4) is determined as:

\[ H(q) = -\frac{1}{k} \sum_{i=1}^{n} \log_2 q_i(x) \text{, bit} \]  \hspace{1cm} (5)

There is a difference between entropy (3) and cross-entropy (4), which is called Kullback-Leibler divergence – the discrepancy between the distribution of \( q \) and \( p \):

\[ D_p(q) = H_p(q) - H(p). \]  \hspace{1cm} (6)

When substituting expressions (4) and (3) into (6), by performing simple mathematical transformations, the divergence of the Kullback-Leibler has the form:

\[ D_p(q) = \sum_{i=1}^{k} p_i(x) \log_2 \frac{p_i(x)}{q_i(x)} \text{, bit} \]  \hspace{1cm} (7)

It should be borne in mind that the functional (7) is not a metric in the space of distributions, since it doesn’t satisfy the axiom of symmetry: \( D_p(q) \neq D_q(p) \).

4. Conclusion
Having considered the main directions in the study of network analysis, it was found that the entropy approach to the assessment of integration processes takes a leading position, since it accumulates a number of developed theories. The entropy measure of information – a universal tool for measuring the state of the network structure to develop rational solutions. The determination of the amount of entropy and its application in the analysis process allows to identify opportunities (importance) of the network elements in supporting the integration processes of the agro-industrial complex. The divergence of Kullback-Leibler, which is the distance (the degree of distinctiveness) between two distributions of a random variable, acquires an important significance here. The presence and capabilities of the entropy interpretations presented in the paper can be considered justified. Its application in solving problems of assessing the structural content of networks and probabilistic behavior of individuals is an integral part of the implementation of a comprehensive solution to the integration processes in the agro-industrial complex.

Acknowledgments
The research was carried out with the financial support Program of the Russian Foundation for Assistance to Small Innovative Enterprises in Science and Technology by the grant "UMNIK" №13138GU/2018, code № 0040353.

References
[1] Khalilova M M 2013 Features and a role of integration processes in agro-industrial complex Issues of structuring the economy Economics and Economics pp 53-56
[2] Powell W W and Grodal S 2005 Networks of Innovators The Oxford Handbook of Innovation
Ed J Fagerberg, D C Mowery, R R Nelson (Oxford; New York: Oxford University Press) pp 56–85

[3] Çetinkaya U Y and Erdil E 2016 Cohesion and Competition of Europe: Innovation Policy from the Perspective of Networks and Entropy Foresight and STI Governance 10 4 7–24 DOI: 10.17323/1995-459X.2016.4.7.24

[4] Asheim B T, Gertler M S 2005 The geography of innovation: Regional innovation systems The Oxford Handbook of Innovation. Eds. J. Fagerberg, D.C. Mowery, R.R. Nelson. Oxford / New York: Oxford University Press pp 291–317

[5] Das T K, Teng B S 2002 Alliance Constellations: A Social Exchange Perspective The Academy of Management Review 27 No 3 pp 445–456.

[6] Sturgeon T, Biesbroeck J V and Gereffi G 2008 Value chains, networks and clusters: Reframing the global automotive industry Journal of Economic Geography 8 (3) 297–321

[7] Gereffi G, Humphrey J and Sturgeon T 2005 The governance of global value chains Review of International Political Economy 12 (1) 78–104

[8] Granovetter M 1973 The Strength of Weak Ties American Journal of Sociology 78 (6) 1360–80

[9] Makrufa Sh Hajirahimova, Aybeniz S Aliyeva 2017 About Big Data Measurement Information Measurement of the Structure State of the Technical System Vestnik UrFU. Economics and Management 14(5) 741-58 DOI: 10.15826/vestnik.2015.14.5.41

[10] Romanova O A, Makarov E V 2015 Theoretical foundations and methodologies for assessing the integration cooperation at the industry market Vestnik UrFU. Economics and management 14(5) 741-58 DOI: 10.15826/vestnik.2015.14.5.41

[11] Kogut B, Zander U 1992 Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology Organization Science 3 (3) 383–97

[12] Allen P M 2001 A complex systems approach to learning in adaptive networks International Journal of Innovation Management 5 (2) 149–80 DOI: 10.1142/s136391960100035x

[13] Dulesov A S, Karandeev D J and Krasnova T G 2017 The evaluation of the correlation between entropy and negentropy in the structure of a technical system MATEC Web Conf 129 1-4 DOI: https://doi.org/10.1051/matecconf/201712903005

[14] Dulesov A S, Eremeeva O S, Karandeev D J and Dulesova N V 2018 Approaches to Information Measurement of the Structure State of Technical Systems IEEE FarEastCon 2018 pp 1-6 DOI: 10.1109/FarEastCon.2018.8602799

[15] Cockshott W P and Cottrell A F 1996 Information and Economics: A Critique of Hayek Research in Political Economy 16 177–302

[16] Popkov Y S, Shvetsov V I, Weidlich W 1998 Settlement formation models with entropy operator The Annals of Regional Sciences, Springer-Verlag 32 267–294

[17] Dulesov A S, Eremeeva O S, Karandeev D J and Krasnova T G 2018 Analytical notes on growth of economic indicators of the enterprise Advances in Economics, Business and Management Research 47 327-32 DOI: https://doi.org/10.2991/iscefec-18.2019.81

[18] Dulesov A S, Karandeev D J and Dulesova N V 2017 Optimal redundancy of radial distribution networks by criteria of reliability and information uncertainty IEEE 3rd International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM) pp 1-4 DOI: 10.1109/ICIEAM.2017.8076467

[19] Dulesov A S, Karandeev D J and Dulesova N V 2017 Reliability analysis of distribution network of mining enterprises electrical power supply based on measure of information uncertainty IOP Conference Series: Earth and Environmental Science (EES) 87 1-6 DOI: https://doi.org/10.1088/1755-1315/87/3/032008

[20] Hartley R V L 1928 Transmission of Information Bell System Tech. J. 7 (3) 535–63

[21] Fazlollah M Reza 1994 An Introduction to Information Theory New York: Dover Publications, Inc.
[23] Johnson E 2018 Anxiety and the Equation: Understanding Boltzmann's Entropy *The MIT Press*
[24] Shannon C E 1948 Mathematical Theory of Communication *Bell System Tech. J.* 27 379-423
[25] Shannon C E 1949 Communication Theory of Secrecy Systems *Bell System Tech. J.* 28 656-715
[26] Dulesov A S, Karandeev D J and Dulesova N V 2018 Determination of the amount of entropy of non-recoverable elements of the technical system *IOP Conf. Series: Materials Science and Engineering, MISTAerospace* 450 1-6 doi:10.1088/1757-899X/450/7/072004
[27] Dulesov A S, Karandeev D J and Dulesova N V 2018 Improving the operation quality of technical systems using information theory models *MATEC Web Conf.* 224 DOI: https://doi.org/10.1051/matecconf/201822404006
[28] Kullback S, Leibler R 1951 On information and sufficiency *Ann. Math. Statist.* 22 (1) 79–86