Process entropy and informational macrodynamics in a ceramic tile plant

Franjo Jović*, Darko Krmpotić**, Alan Jović***
*Department for Computer Engineering and Informatics,
Faculty of Electrical Engineering, University J.J. Strossmayer in Osijek,
31000 Osijek, kneza Trpimira 2b Tel: 031-224-615 Fax: 031-224-605 E-mail: jovic@etfos.hr
**KIO-Keramika d.o.o. Orahovica,
***Faculty of Electrical Engineering and Computing, University of Zagreb

Abstract - Product quality concept in a ceramic tile plant is presented. Entropy and information content of measurement data is calculated. Several entropy concepts have been elaborated: Shannon entropy, 1D and 2D Carnap entropy. Because of plant informational macrodynamics the dominance of 1D Carnap entropy of process data sources has been argued on measurement data of the kiln temperature profiles and press control parameter data. Obtained process entropies have been compared to normalized process data distributions such as Gaussian and uniform distribution.

I. INTRODUCTION

Ceramic plant behavior and respective quality of its final product highly depend on the overall plant organization, process technology state and production discipline. In order to study the regularities in the plant results and to enable detection of its quality critical points a series of measurements has been made and its parameter studied. Among the most serious analysis is the analysis of process entropies. In order to fulfill the plant goal the information system theory has been proposed [1] that unifies information description of different process information sources and information formalism for building the system regularities. Such system regularities are represented by informational macrodynamics.

The self-information of a random event or a random message is a term coined by C. E. Shannon who defined it to be "minus logarithm of the probability of a random event". The Shannon "entropy" of the stochastic source that generated the event is the expectation of the self-information. Shannon discovered that the entropy of a stochastic source has a clear and important physical meaning: the smallest number of bits to represent the event from the stochastic source.

In the broadest sense the source distribution is completely known, or it is known to belong to a parameterized family of probability distributions, or the source distribution is known to be stationary and ergodic, but no other information is available [2]. A question arises: can we find an appropriate and universal way to estimate the probability measure that describes the generation of information at the information source?

For short data acquisition series the probability being approximated with relative frequencies cannot be defined. Sometimes no *a priori* information about the underlying statistics of the source is available. Even worse, the majority of real production processes possess strict and controlled number of source values with superimposed unknown noise. On the other hand processes exhibit standard information patterns representing expected states of the plant. These patterns do not follow any distribution, violating the possibility of relevant information measurement.

Informational description of the plant cannot strictly follow from Shannon's information entropy concept because of evident deterministic structure that results in data uniformity disruptions. Thus several other entropies have been studied and elaborated in the article in reference to the plant data dynamics. Entropy results have been presented and compared.

II. PROCESS ENTROPIES

Information system theory as proposed by Vladimir S. Lerner [1] is conceived to build a bridge between the mathematical systemic formalism and information technologies dealing with transformation of information. The aim of such procedure is to obtain system models that reveal information laws and specific object codes and patterns. General evaluation of informational aspect of stochastic data is enabled by application of Kolmogorov entropy [3] that measures the extent of chaos in a system.

According to Shannon's interpretation, system uncertainty is closely related to system information content we possess [2]. Let us designate with $I(l,T)$ the information amount obtained by following up system trajectories in time interval T with precision l. Basic cell of the system phase space with dimension m possesses the volume l^m and observation time interval $N\tau = T$. Designating with $p_{i0,i1\ldots iN}$ complex probability of the observed system $x(t)$ to be found in $t=0$ in cell i_0 , in $t=\tau$ in cell i_1,..., and in $t = N\tau$ in cell i_N, the informational content equals to

$$I(l,T) = -\sum_{i0,i1\ldots iN} p_{i0,i1\ldots iN} ldp_{i0,i1\ldots iN} \qquad (1).$$

Kolmogorov entropy is given as the limit value of

$$K = \lim_{l\to 0}(\lim_{T\to\infty}\frac{1}{T}I(l,T)) \qquad (2).$$

Kolmogorov entropy thus measures information amount necessary for precise determination of system trajectory in phase space, or the loss of information in the system starting from its initial point. For system states with equal probabilities $p_{i0,i1\ldots iN} = \frac{1}{N}$, and $K \approx N\frac{1}{N}ldN = ldN$.

The generalization of K entropy can be done by introducing so called information of the order q as [4]

$$I_q(l,T) = \frac{1}{1-q} ld \sum_i p_i^q \qquad (3).$$

Information content is information of the first order. K-entropy of the order q is defined as in [2]:

$$K_q = \lim_{l \to 0}(\lim_{T \to \infty} \frac{1}{T} I_q(l,T)) \qquad (4).$$

K_0 is topologic entropy, K_1 is information entropy, K_2 is correlation entropy etc. Permitted are also non-integer values of parameter q.

Topologic entropy excludes the notion of probability and treats the abstract measure of entropy as introduced by philosopher Rudolf Carnap [5]. By constructing a time-dynamic two dimensional Voronoi diagram using Voronoi cell generators with coordinates of value and change of value, entropy becomes a function of the cell areas. Accepting the knowledge of the final (teleonomic, purposeful) state of the system at each stage of its trajectory the 2D Carnap or teleonomic entropy is used to describe changes in any end-directed system. Actually this term can be justified only at the price of system state enumeration follow-up.

We will introduce a simplified approach to the topologic entropy concept by reducing the dynamic dimension of the process, i.e. excluding system dynamics and observing the system as a series of its values only. This approach significantly reduces calculation burden when defining specific event space distribution but introduces the problem of multiple identical state values that was hidden in the original Carnap concept and in its derivates as well.

III. A REVISITED CARNAP ENTROPY CONCEPT

Rudolf Carnap (1956) introduced an n-dimensional system space using its n system variables within their theoretic limits R^{μ}, μ being at maximum equal to n. Each system variable u_i is given within its minimum and maximum values of its space ϕ_i. A two, three, or more – dimensional spaces are produced with respective Voronoi diagrams. The relevant occupation space e_j is defined for each measurement point $b_j(u_{i1}, u_{i2})$ according to minimum distance criterion from neighboring points.

Because of the possibility to relate the relative ratio of each occupation space and theoretic limit, the dual logarithm of the relation is named Carnap entropy:

$$I_C = -\sum_j \frac{e_j}{R^{\mu}} ld \frac{e_j}{R^{\mu}} \quad (\text{ bit }) \qquad (5).$$

The time-space trajectory of the system produces a dynamic Voronoi diagram that can be traced and compared to the desired, teleonomic system trajectory. Thus Carnap entropy can measure system entropy of teleonomic systems.

A problem not specified in the Carnap concept is the case of equal measurement points and the case of extreme reduction of the dimensions to only one.

It is important to stress out that multiple point evaluation exists. The reason for multiple point evaluation exists in technology praxis where strict control of particular process measurement is necessary for product quality enforcement. On the other hand one-dimensionality of the Carnap entropy is forced whenever overall plant data are examined whose mutual difference values do not indicate any system dynamics or just do not carry explainable information content.

IV. ONE-DIMENSIONAL CARNAP ENTROPY FOR PERIODIC, UNIFORM, AND GAUSSIAN DISTRIBUTION

Theorem 1: Carnap based entropy is invariant to event occurrence instant.
Proof: The change of element ordering in the relation (5) does not produce any change in the sum value.
Corrolary 1: One-dimensional Carnap entropy as stated in relation (5) is applicable also to time independent data series.

Periodic scanning N times of a constant process value L should yield a negligible one-dimensional Carnap entropy because the space that occupies first measurement is total space L with $\frac{L}{L} ld \frac{L}{L} = 0$ and all the other data occupy a negligible space with $\frac{0}{L} ld \frac{0}{L} = 0$.

However, reserving for $N-1$ measurement the $\frac{1}{N-1}$ part of the space can yield the entropy in the form that is independent of L that is $\lim I_{CL} = \lim_{N \to \infty}\left(-\frac{N-2}{N-1} ld \frac{N-2}{N-1} - ld \frac{1}{N-1}\right) = \infty$.

Thus for N=25 Carnap entropy equals to 0.0614 + 4.5849 bit, for N=1000 it equals 0.00144 + 2.9996 bit, and for N=100.000 it equals 0.0000043 + 16.6096 bit which is not a satisfying result.

Therefore space of equal measurement values should not be subdivided among equal measurement points. It should be taken as a unique however small space e_j and its content multiplied by the number of equal measurements. This statement is valid for Carnap entropy of any dimension.

IV.I. 1D UNIFORMLY DISTRIBUTED CARNAP ENTROPY

Let us suppose that a uniformly distributed one-dimensional data exhibit total number of occurrences N put into $k-1$ equal increment classes with incremental differences Δ_i and also that there is defined an initial class w_0. Then a total space *L* that can be expressed as

$$L = w_0 + (k-1)\Delta_i \quad (6).$$

Let us suppose that w_i represents the measured value of the i^{th} class. Supposing the initial value of the uniform distribution possessing value *x*, $x \overset{def}{=} <0, \mu>$, *x* being a left flank of the class w_0, and knowing that the mean measurement value equals to $\mu = \frac{L}{N}$, then the value Δy between *x* and mean value equals to $\Delta y = \mu - x$ and the class steps uniformly increase from *x* to $x + 2\Delta y$ with class increments $\Delta_i = \frac{2\Delta y}{k}$. Thus the measured value of the class *i* equals to

$$w_i = x + i\left(\frac{2\Delta y}{k}\right),\ i = 1,2,\ldots,k \quad (7).$$

Possessing $\frac{N}{k} = \lambda$ events in each class total 1D Carnap entropy of the uniform distribution is

$$C_U = -\lambda\left(\frac{x}{L} ld \frac{x}{L} + (k-1)\frac{\Delta_i}{L} ld \frac{\Delta_i}{L}\right) \quad (8).$$

Usually the number of classes *k* is approximated with $\sqrt{N}$, and relation (8) depends only on the value x.

IV.2. 1D GAUSSIAN DISTRIBUTED CARNAP ENTROPY

Gaussian distribution involves measured values distributed into classes with different number of events with values distributed according to given parameters (μ, σ^2) of mean value and variance. With Δ as interval between measurements, *N* as total number of measurements, total occurrence space *L* is given by $L = \max_{\mu,\sigma}(N)$. Occurrence classes are usually of the width $c \cong \frac{\sigma}{2}$. Left flank of the distribution can be calculated from the binominal distribution as [6]

$$f_{ti} = Np_i = 1 \quad (9),$$

where f_{ti} is expectation of one event in the minimum value class. With known N, the correspondent deviation from mean value for unity normal distribution can be calculated from Gaussian unity distribution as $\varphi\left(\frac{1}{N}\right)$. For $N = 100$ the distribution is between $(-3.54\sigma, 3.54\sigma)$ and data can be subdivided into 14 classes. Thus one-dimensional Carnap entropy of the Gaussian measurement starts from $\mu - \varphi\left(\frac{1}{N}\right)$. The number of classes is $k = \left|\frac{2c\varphi\left(\frac{1}{N}\right)}{\sigma}\right|$. Therefore total 1D Carnap entropy of the Gaussian distribution is

$$C_G = -\frac{f_{t1}(L - 2\varphi\left(\frac{1}{N}\right))}{L} ld \frac{f_{t1}(L - 2\varphi\left(\frac{1}{N}\right))}{L} - \\ -\sum_{i=2}^{k} \frac{f_{t1}c}{L} ld \frac{f_{t1}c}{L} + err \quad (10).$$

The error term has to be added because of volatility of distribution that can change upper and lower flank of the actual distribution. The inclusion of actual data for f_{ti} and with $f_{t1} = 1$ gives

$$C_G = -\frac{L - 2\varphi\left(\frac{1}{N}\right)}{L} ld \frac{L - 2\varphi\left(\frac{1}{N}\right)}{L} - \\ -\sum_{i=2}^{k} \frac{Nc^2\varphi(u_i)}{\sigma L} ld \frac{Nc^2\varphi(u_i)}{\sigma L} + err \quad (11).$$

V. PROCESS MEASUREMENT DATA AND RELEVANT ENTROPY CALCULATIONS

Data have been taken from two-stage kiln temperature sensors on the daily basis from August to October 2008 in the KIO KERAMIKA tile factory in Orahovica, Croatia. A separate program that calculates Shannon and 1D and 2D Carnap entropies has been prepared in Java programming language. Calculations are given in Table 1.

The following 31 signals, Signal 10 to Signal 40, have been taken across five tile press machines measured on the daily shift basis and relevant entropies have been calculated as in Table 1. Data are given in Table 2.

VI. DISCUSSION AND CONCLUSION

Data on process measurements are required for further analyses, modeling, decision and intervention. Common issue in process analyses and modeling are product quality while retaining adequate plant productivity. In KIO KERAMIKA ceramic tile plant high productivity is enabled through complete press, design prefabrication, and kiln automation as well as by robotic platform application in the internal tile transport. Still high production machines

Table 1: Entropy calculations of 90-day temperature profiles in the two-stage kiln in KIO KERAMIKA tile factory, Orahovica, Croatia

Signal name	Temperature lower bound, ˚C	Temperature upper bound ,˚C	Shannon entropy, bit	1D Carnap entropy, bit	2D Carnap entropy, bit
Signal 44	1106	1166	none*	17,0464	5,82848
Signal 45	1119	1151	2,89715	20,4124	6,96516
Signal 46	1106	1166	none	16,8702	5,44023
Signal 47	1106	1166	2,9183	19,6232	6,66346
Signal 48	1107	1169	none	15,7453	5,5408
Signal 49	1114	1150	2,74091	18,5061	5,98589
Signal 50	1107	1170	none	16,0646	5,584
Signal 51	1114	1150	2,7879	18,7413	6,5268

* one or more empty data classes

Table 2. Entropy calculations for 31 signals, Signal 10 to Signal 40, across five tile press machines for daily shift quality control

Press name	Data lower bound	Data upper bound	Shannon entropy, bit	1D Carnap entropy, bit	2D Carnap entropy, bit
Press 5	-0,1	50,1	none*	5,01798	3,29666
Press 6	0	60,1	none	3,16758	3,05867
Press 8	0	48	none	4,59378	3,48107
Press 9	0	50,1	none	3,08708	3,37782
Press 10	-0,1	47,1	none	4,76176	3,30893

*- one or more empty data classes

require adequate workers' support and logistics. Therefore 55 signal sources in the process are measured and taken from the process by quality control division of the plant mostly on the daily shift basis and independent of process control manipulations. As referred by cited authors [1, 2, and 3] and known from Shannon's information theory, the quantity of information can be measured by data entropy. Stochastic information source with lower entropy contentthereby contains more information, because of its lower internal data volatility. Thus process information source with high volatility can be indicated as a source of quality problems in the production.

Data time series taken during three months from the ceramic kiln and put into Shannon based procedure have shown poor ability for usual entropy analysis because of zero values in their distribution function that are the basis for entropy calculations, Table 1. Moreover data taken across presses cannot completely be used for such analytical purposes, Table 2. Therefore a new concept of one-dimensional Carnap entropy has been proposed that enables simple and fast process entropy calculations. This entropy covers complete measuring range by dividing it into subspaces of each particular measurement value.

Authors propose that multiple data occurrences in the Carnap theory should be solved by addition of particular entropy contributions. The obtained results with 1D Carnap entropy for press quality indicate the difference in production quality for press 6 and 9 compared to press 5, 8 and 10. Carnap entropy is usually conceived as two-dimensional entropy of the signal state dynamics [4, and 5]. This entropy is calculated for process signals as well, and given in Tables 1 and 2, but for process across signal data it is not interpretable. The procedure is also very time-consuming, taking approximately ten minutes for each 31-data row in Table 2 calculations on a 2 GHz machine with 512 MB memory and Intel processor. Algorithm calculation optimization was not applied, although around 3 billion operations are required for Voronoi diagram calculation in the particular case.

Further work is required to connect calculated entropies for the whole production process with product quality models and with worker's motivation profiles into a meaningful and harmonic unity [1].

ACKNOWLEDGEMENTS

Authors are thankful to Mrs. Verica Hocenski M.Sc., quality control manager of the KIO KERAMIKA d.o.o., for support in data preparation.

REFERENCES

[1] V. S. Lerner, "Information System Theory and Informational Macrodynamics: Review of the Main Results", *IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and reviews,* vol. 37, no.6, p. 1050-1066, 2007.

[2] A. D. Wyner, J. Ziv and A. J. Wyner, "On the Role of Pattern Matching in Information Theory", *IEEE Transactions on Information Theory*, vol. 44, no. 6, p. 2045-2078, 1998.

[3] A. N. Kolmogorov, *Information Theory and Theory of Algorithms*. Nauka, Moscow, 1987.

[4] P. Grassberger and I. Proccacia, "Dimensions and Entropies of Strange Attractors From a Fluctuating Dynamics Approach", *Physica D*, vol. 13, issue 1-2, p. 34-54, 1984.

[5] A. Pudmetzky, "Teleonomic Entropy Measuring the Phase-Space of End-Directed System", *Applied Mathematics and Computation*, vol. 162, no. 2, p. 695-705, 2005.

[6] I. Pavlić, *Statistička teorija i primjena*, Tehnička knjiga, Zagreb, 1988.