

D. Tewari, G.T. Heydt

Author Affiliation: Power Systems Engineering Research Center, Arizona State University, Tempe, AZ, USA.

Abstract: A novel method of signal analysis, symbolic dynamics, is employed to model highly varying electrical loads. The method is based on a discretized form of the load current signal. Symbolic dynamics is based on processing similar to natural languages. The method is applied to ac electric arc furnace modeling.

Keywords: Symbolic dynamics, load modeling, power conditioning, power quality.

Power Quality Representation: The general process of signal analysis presents a basic quandary between the information lost in processing versus the ease and convenience of obtaining a measure of the signal. As an example, distorted 60 Hz voltage waves in primary distribution systems may be analyzed using the total harmonic distortion (THD). The value of THD_v is easily calculated, but may not fully characterize the signal $v(t)$ for a given application. In fact, in some cases, the total harmonic distortion may not even exist (e.g., there is no fundamental component; the signal is aperiodic and does not possess harmonics) [1].

There are several mathematical approaches that utilize the *state evolution method* as a basis of analysis of electric power quality. In this concept, a time series of values $v(t_1), v(t_2), \dots$ is used to represent $v(t)$. One might argue that any discrete analysis uses this approach. Markov chain models are a good example of the state evolution approach [2], [3].

An Introduction to Symbolic Dynamics: The term *symbolic dynamics* refers to the state evolutionary mathematical approach in which signal discretized states are denoted by symbols [4], [5]. Under such a notation, the signal $v(t)$ becomes a time sequence, for example:

A	B	C	D	...
---	---	---	---	-----

The elements A, B, C, D, ... are termed symbols and they may represent voltages, currents, their derivatives, or other physical parameters. Certain groups of (or sometimes single) symbols are termed *words*. It is possible to scan through a large volume of time series data to identify the recurring *words* by the process of identifying recurrent sequences. It is also possible to identify longer strings of words and these strings are termed *sentences*. A *dictionary* is a collection of words formed from a signal. A dictionary can be a multidimensional array, containing words from different signals, for example a voltage signal and a current signal. The number of times a word appears in a dictionary is called the *frequency* of that particular word, and the *fractional occurrence* is the ratio of the frequency of a word to the total number of words in that dictionary. The notation f_{xi} denotes the fractional occurrence of the i th word in dictionary X.

There are advantages of Fourier-based analyses that are, unfortunately, absent from the symbolic dynamic approach. Consider the time shift property of Fourier transforms and Fourier series: this property allows one to shift the signal $v(t)$ in the time domain and obtain very similar Fourier series (in the case of periodic signals) and very similar Fourier transforms in general. A second consideration relates to magnitude scaling of time domain data. In the frequency domain, two signals are easily compared at several frequencies (or a whole band of frequencies) by simple magnitude scaling. Once the time signal is coded as a symbolic sequence, scaling is not convenient. The solution to this difficulty is to scale the signal participants in the time domain before moving to the symbolic dynamic domain.

The foregoing difficulties come with concomitant advantages: one may compare any two signals, periodic or aperiodic; there is no dependence on sinusoidal waveshapes; and noise gives less trouble.

Symbolic Dynamic Dictionary Formation: A dictionary is a two-dimensional array for a scalar signal. This array can be filled by considering all the one-symbol words, one at a time, then two consecu-

tive symbols at a time and so on, until n_w symbols at a time, where n_w is the maximum word length defined for the dictionary. The dictionary also contains a parallel vector that is filled with the fractional occurrences of each word. Figure 1 shows the process

$$\chi = [Q, W, E, R, T, Y, U, I, O, P, Q, H, \dots, E, G, H].$$

To compare two signals, a reference dictionary is formed from one signal and a dictionary is formed from a test signal. The two dictionaries are compared row by row and a mini dictionary is formed from the common words. This mini dictionary also consists of two columns indicating the fractional occurrences of each word in each dictionary. The process is shown in Figure 2 for signals X and Y.

The common signal index (CSI) can be calculated by using the formula

$$CSI = \sum (f_{xi} * f_{yi}) / (f_{xi} + f_{yi})$$

where f_{xi} is the fractional occurrence of the i th common word in the χ dictionary, f_{yi} is the fractional occurrence of the i th common word in the γ dictionary, and CSI_{max} is 0.5 (if the two signals are exactly identi-

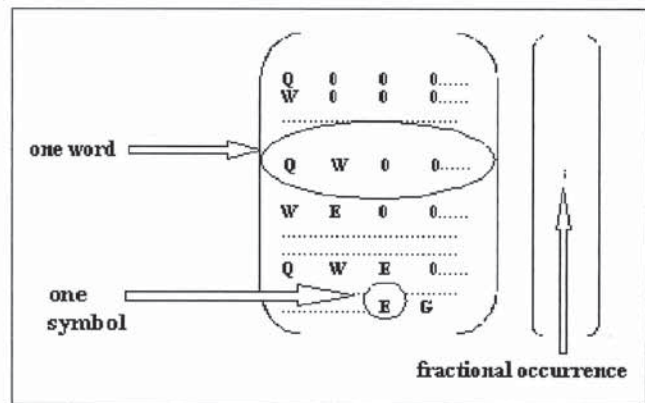


Figure 1. Forming a symbolic dynamic dictionary

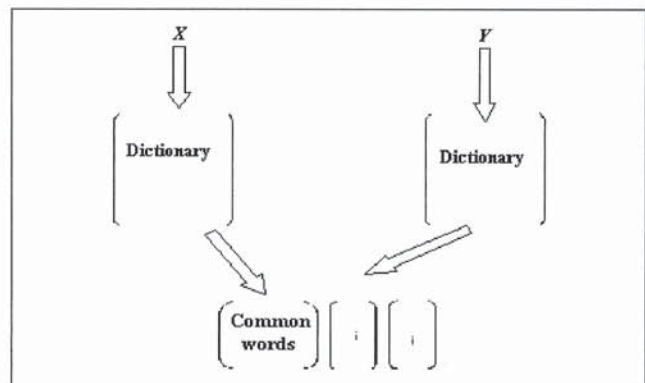


Figure 2. Comparing two signals

Table 1. Results for test 1					
Index		Actual Load Current	Predicted Load Current		
			Trial 1	Trial 2	Trial 3
RMS*		2.06	2.06	2.16	2.16
% RMS Deviation	Historical		4.22	8.84	9.17
	Actual		0.35	4.79	5.11
CSI		0.5	0.44	0.47	0.46
KS significance level α		1	0	0	0
* historical RMS value: 1.98 p.u.					

* historical RMS value: 1.98 p.u.

Table 2. Error in the average power for test 2

Trial	Absolute % Error in Average Power of the Predicted EAF Load Current		
	Phase A	Phase B	Phase C
1	8.13	14.71	8.73
2	6.37	13.01	0.45
3	0.19	6.24	10.18
4	0.19	6.24	10.99
5	3.98	14.71	7.22
6	10.09	13.01	5.87
7	3.45	14.71	11.12
8	9.09	6.24	10.35
9	4.37	6.24	8.73
10	0.66	13.01	0.45
Expected error (%)	4.65	10.81	7.41

cal). The lower the value of CSI, the lesser the degree of similarity between the two signals.

A second signal comparison technique is known as the Kolomogorov-Smirnov (KS) method. The KS method involves identifying empirical distribution functions for the two populations and then assessing whether there are any differences between the two distributions [6]. In the KS test, one obtains the significance level α , which is a measure of similarity in the two populations.

Symbolic Dynamics and Load Modeling: Attention turns to the forecasting or prediction of future values of a signal. For this purpose, a dictionary is formed utilizing the instantaneous values in a signal in a time series. The prediction is accomplished by matching the most recent time series measurement with dictionary entries. Entries with high fractional occurrence are selected as predictions.

The concept outlined above was applied to an ac electric arc furnace (EAF). The data used for this purpose have been recorded from a 34.5 kV, 50 MVA ac electric arc furnace under normal operating conditions. The arc furnace current in different phases is forecasted using the algorithms, assuming that the supply voltage is fixed amplitude ac. The results obtained are compared with the actual measured EAF current. The comparison is done for RMS current, CSI, and KS significance level α . The data obtained directly from an EAF have been measured with a sampling frequency of 10 kHz. The number of measurements is 500 points in three cycles. Two tests will serve to typify results (other tests are reported in [7]):

- Test 1: In this test the historical data set consists of 36 cycles of current and predicted data are for 24 cycles. The symbolic dynamic current is compared with historical time series measurements in Table 1.
- Test 2: The aim of this test is to verify the proposed model from the point of view of active power utilization. The value of voltage was recorded with the same sampling frequency as the current. Using symbolic dynamics, the future current value was predicted. Assuming the voltage to be constant amplitude, the average power taken by the load was calculated. Comparison of the predicted and actual average power for ten different trials has been reported in Table 2 for each phase.

Discussion: For load modeling applications, the proposed algorithm is executed offline. The symbolic dynamic dictionary is not further updated after formation. From the point of view of comparison, the RMS value of the EAF load current is an important parameter. The CSI and KS statistics are the measure of similarity between the two signals. These two indices are related to the waveshape of the actual and modeled signal. In load modeling, the detailed waveshape of the predicted signal is often of lesser importance. The load power depends mainly on the RMS value of the current and its phase. The application shown is for

load forecasting using symbolic dynamics. There is a potential for the use of high-speed processing of historical data by this method for power conditioning: in this application, symbolic dynamics is used to obtain a correction current which, when injected into the load circuit, causes the resultant load current to be nearly sinusoidal [7].

Conclusions: An innovative concept from signal processing, symbolic dynamics, has been applied to the characterization of a highly varying load current for an electric power load. The method entails coding the signal in symbols that are discrete levels of the signal. The symbols are processed as if they were in a natural language. The method is applied to load modeling and symbolic dynamics appears to give reasonable accuracy of RMS load current and active power demand for the case of an ac electric arc furnace. In the area of load current smoothing, the method offers potential as an active filter.

Acknowledgments: The authors are grateful for the contributions to this work by Dr. J. Rico, The Arizona Public Service Co., and the Power System Engineering Research Center (PSERC).

References:

- [1] G.T. Heydt and W.T. Jewell, "Pitfalls of electric power quality indices," *IEEE Trans. Power Delivery*, vol. 13, pp. 570-578, Apr. 1998.
 - [2] A.M. Stankovic and E.A. Marengo, "A dynamic characterization of power system harmonics using Markov chains," *IEEE Trans. Power Systems*, vol. 13, pp. 442-448, May 1998.
 - [3] S. Arora, D. Salem, and G.T. Heydt, "Innovative concepts in power system load modeling," in *Proc. North American Power Symp.*, Waterloo, Ontario, Canada, Oct. 2000, pp. 15-47.
 - [4] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*. Cambridge, U.K.: Cambridge University Press, 1995.
 - [5] C.W. Wu and L.O. Chua, "Symbolic dynamics of piecewise-linear maps," *IEEE Trans. Circuits Syst. II: Analog and Digital Signal Processing*, vol. 41, pp. 420-424, June 1994.
 - [6] M. Hollander and D.A. Wolfe, *Nonparametric Statistical Methods*. New York: Wiley, 1999.
 - [7] D. Tewari, "Symbolic dynamic models for highly varying power system loads," M.S.E.E. Thesis, Arizona State Univ., May 2002.
- Copyright Statement:** ISSN 0282-1724/02/\$17.00 © 2002 IEEE. Manuscript received 28 November 2001; revised 23 May 2002. This paper is published herein in its entirety.

2002 Conference on Power Systems and Communication Systems Infrastructures for the Future

23-27 September, Beijing, China

The Conference on Power Systems and Communication Systems Infrastructures for the Future will be held in Beijing, China on 23-27 September 2002. The conference brings together researchers, manufacturers of infrastructure hardware and software systems, and institutions involved in emergency management at national and international levels, in order to exchange views on what should be the direction of developments in these critical disciplines.

The conference Web site, <http://www.cris.vt.edu>, provides details about the conference organization and facilities. For more information, contact A.G. Phadke, Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24061 USA, +1 540 231 7029, fax +1 540 231 4303, e-mail aphadke@vt.edu.