Neural network analysis of electrodynamic activity of yeast cells around 1 kHz

R Janca
Czech Technical University in Prague, Faculty of Electrical Engineering, Prague, Czech Republic
E-mail: jancarad@fel.cvut.cz

Abstract. This paper deals with data analysis of electrodynamic activity of two mutants of yeast cells, cell cycle of which is synchronized and non-synchronized, respectively. We used data already published by Jelinek et al. and treat them with data mining method based on the multilayer neural network. Intersection of data mining and statistical distribution of the noise shows significant difference between synchronized and non-synchronized yeasts not only in total power, but also discrete frequencies.

1. Introduction
Electromagnetic activity and sensitivity of single cells has been extensively studied in the past [1]. Despite a number of indirect observations and theoretical models [2], electromagnetic activity of single cell has been measured only with limited success. However, advances in nanotechnology have led to reopening of this issue [3].

In recent paper by Jelinek et al., the evidence was shown for the electrical activity of yeast cells Saccharomyces cerevisiae in the frequency range 0.4 – 1.6 kHz [4]. However, this conclusion was based mainly on the mean electrical power in entire frequency range measured for two populations of cells in different phase of the cell cycle. One may have reservations about this method since it may be strongly influenced by the change in bio-chemical and electrical properties of the cells and the buffer due to metabolic activity of cells in certain phase of the cycle. In order to resolve whether this objection is valid, we employed artificial neural network in the analysis of data presented in [4]. In this paper we present results of the assay which revealed significant difference between two cellular cultures at individual frequencies. This difference, analysis proved, may not be attributed to technical artefact or chemical reason.

2. Materials and methods
We used data kindly provided by authors of Jelinek et al. [4]. In following we present a brief summary of measurement system and procedure. Detailed description is given in [4].
2.1. Yeast cells

The β-tubulin tub2-401 mutants of yeast cells Saccharomyces cerevisiae (strain CUY67 Mata tub2-401 ura3-52 ade2-101) were used for the measurement. The mutant cells were cultivated under the restrictive temperature of 14°C when microtubules cannot polymerize. Most cells interrupt their cell cycle before entering into M-phase, because mitotic spindle cannot be assembled. The microtubules begin to polymerize after heating of the cells above the permissive temperature of 25°C, so the majority of cells assemble the spindle and start the mitosis synchronously (synchronized cells). The part of the mutant cells was cultivated at the temperature of 30°C. These cells were not synchronized (non-synchronized cells) and they were used as a reference.

2.2. Laboratory equipment

The measurement system is depicted in Figure 1. The important part of the system is the sensor followed by transform and power amplifiers placed in a triple electromagnetic shielded box which is housed inside a temperature stabilized chamber. The sensor is a sharp platinum wire with tip that is few micrometres above the bottom of the cuvette. Space between the bottom of the cuvette and the tip is comparable with the size of the yeast cell. Cells settle on the bottom of the cuvette, so the sensor can contact only one cell (see schematically in Figure 2). But a large part of the platinum wire is still immersed in the suspension, so the impedance of the suspension contributes to the background noise. The sensor is followed by the transform amplifier that is connected before the power amplifier. The gain of the amplifiers is around 110 dB at 1 kHz. Both amplifiers are supplied by batteries that are placed inside the shielded box. This helps to minimize spurious signals. The thermal box stabilizes temperature inside at 28±0.5°C with minimal hysteresis. Output of the power amplifier is connected to the spectrum analyser Agilent E4448A. Theoretical sensitivity of the system is $10^{-18}$÷$10^{-19}$ W, that is the thermal noise level.

2.3. Measurement

A suspension of the cells has optical density 4.5 (OD 600), which represents a concentration of about 2·10^8 cells per milliliter. The synchronized cells were cultivated under the 14°C and non-synchronized around 30°C. Before the measurement, a cuvette with the suspension was placed in 28°C water bath for 3 minutes. The sensor cuvette was then filled with the shaken suspension; volume was approximately 0.06 ml. The measurement of the electrical oscillation started immediately after filling. The spectral analyser measured in frequency range 400÷1601 Hz. Resolution and video bandwidths were set to 1 Hz, reference level was -80 dBm. Whole frequency band was scanned in two sweeps of 601 Hz; both sweeps and storing of data to the PC lasted 6 second. The 400 spectra were recorded during 40 minutes of the measurement.

3. Data analysis

The 49 sets of cells culture, 23 of which were synchronized and 26 were non-synchronized (as a reference), were measured during the experiment. Spectrum of the signal was recorded over time for
40 minutes. The spectrum had 1/f-noise character that is added to transfer function of the amplifiers [5]. The shape of the transfer function was estimated as a time-average spectra of a set, see Figure 3. The shape was depending on the impedance of measured suspension and it could be well approximated by polynomial of 3rd degree which corresponded with electrical stimulation of the amplifiers, an example is given in Figure 4. The spectra were independent on the transfer function after subtraction of its shape.

Figure 3. An example of the spectra and the transfer function estimated as a time-average spectra.

Figure 4. Transfer function estimated as a time-average spectra and its polynomial fitting.

3.1. Power analysis

Figure 5 shows normalized mean spectral power of all of synchronized and non-synchronized sets in time. Non-synchronized cells have uniform trend of the power emission in time. Synchronized cells have increased power between 5th and 20th minutes that correspond with the metaphase and anaphase. Average total power of the synchronized set was 7 % greater that of the non-synchronized. However, the growth and metabolic activity of yeasts during the mitotic process changes parameters of the suspension such as the impedance. The power of the thermal noise strongly depends on the impedance, so difference between sets do not have to be caused by the electrodynamic activity only.

Figure 5. The normalized average spectral power of all cells sets in time. The power of the synchronized cells is increased between 5th and 20th minutes.

3.2. Frequencies analyses

The transfer function and thermal noise are directly dependent on the impedance of the suspension. A frequency analysis minimizes the mentioned dependency. The measured data have noise character
without high spectral lines. Standard methods cannot find significant difference between discrete frequencies of the synchronized and non-synchronized sets. Therefore a method of the data mining using a neural network classifier was employed.

4. Data mining using a neural network

The method is based on following hypothesis:

- A multilayer neural network correctly classifies the spectra into two classes (synchronized, non-synchronized). The spectra do not contain information about an average spectral power. If previous claims are true, the information about difference in the spectra exists. Then weights of network’s inputs show on important frequencies.

4.1. Multilayer neural network

A multilayer neural network (MLNN) is a complex classifier composed of simple classifiers called neurons; more details are in [6]. The McCulloch-Pitts neuron model is depicted in Figure 6. An input vector \( \mathbf{x} \) of data is multiplied by a vector of the weight \( \mathbf{w} \). A sum of multiplied input values and thresholds \( \theta \) is an argument of the transfer-function of the network. An output \( y \) of the neuron is a return value of the transfer-function \( f() \). The index \( j \) is a neuron number; the variable \( n \) represents number of the input parameters. A first and second hidden layer use logarithmic-sigmoidal transfer function. The output layer uses linear function. The scheme of the topology1202-1801-300-2 of the MLNN is shown in Figure 7. A numerical designation of topology represents number of neurons in layers. The topology and transfer-function were chosen as the best of few variants [5].

![Figure 6. The McCulloch-Pitts model of the neuron.](image)

This neural network allows a transformation from 1202 to 2-dimensional space. The MLNN is realized in Matlab 7.9 using Neural Network Toolbox. A training algorithm adjusts neurons parameters (weights and thresholds) to minimize classification error on a training data set. The \( \text{learnngdm} \) training-algorithm was used; it is a gradient descent method with momentum weight and bias learning function which prevents over-learning [6]. The training data contained input and target data. Trained MLNN is able to classify input test vectors of data into two classes (synchronized or non-synchronized). In other words, the neural network finds an optimal function that transforms data to known outcomes. The function can be very complex and cannot be determined through analytical methods. The network can be trained as a classifier which divides testing data into classes. A complex decision making strategy is composed from simple neuron functions.

4.2. Parameterizing of the data

Measured data had to be modified before using the MLNN. The polynomial shapes of amplifier’s transfer function were estimated and subtracted from spectra. Spectral lines of each discrete frequency
were normalized by maximal values of spectral lines throughout all sets. Duration of oscillations is expected less than 30 second. This time corresponded with 5 following measured spectra. Therefore 5 point moving average filter was used for highlight of oscillations and suppression of random high spectral lines in a frequency. The filtered spectra were powered to two in order to improve difference between high and low spectral lines. The parameterized spectra were used as input vectors, input of each corresponded with the frequencies. Additional parameters of the input vectors may be coefficients of the polynomial approximation of the polynomial shapes of amplifier’s transfer function. These coefficients represented the average spectral powers. Using of additional coefficients only verified dependency classification on the average power. Each input vector had a target vector that contained two values representing class of the cells set. The synchronized set represented vector [+1, -1], the non-synchronized [-1, +1].

4.3. Data mining

The 2/3 of parameterized data was randomly separated to the training sets and the 1/3 to the test sets. The data were without additional power parameters. The MLLN was trained by the train sets with the targets vector representing the classes. A validating was proved by the test sets, but the output vector did not contain discrete values. The Equation (1) describes a rule of the assignment to the class. The vectors of the classes of the test sets are \( \text{X} \); the \( \text{Y} \) is output vector of the trained net. If difference of the output vector \( \text{Y} \) has same sign like difference of the classes’ vector \( \text{X} \), the classification is correct (that is true); and conversely.

\[
\text{sign}(\text{X}(1) - \text{X}(2)) = \begin{cases} 
\text{sign}(\text{Y}(1) - \text{Y}(2)) & \text{true} \\
- \text{sign}(\text{Y}(1) - \text{Y}(2)) & \text{false}
\end{cases}
\]  

(1)

A percentile of the true classification represented an ability of learning of the network. Inputs weights of the trained net have described strength of the input parameter for the classification. The strength \( W \) of the input was calculated as average value of the weights of the neurons in the first hidden layer; the \( m \) is a number of the neurons, the \( w \) is weights and the index \( n \) represents the input parameter; see equation (2).

\[
W_n = \frac{1}{m} \sum_{i=1}^{m} w_{ni}
\]  

(2)

The similar method is used to pruning of a net that decrease number of neurons in the net; neurons with small weight are removed [7].

The described process was made 20 times for minimizing random errors. Greater number of repetitions does not significant improve the results. The final results were average of the cross validation.

5. Results

The neural network trained by data with the information about mean spectral power had a validation of the classification around 90 % which corresponded with the power analyses. Data without power parameters had less validation, but the value around 75 % was accepted. Importance of inputs parameters for classification success is represented by input weights. The input weights correspond with frequencies, see figure 8. The high weight shows on the significant frequencies that are the most important for data division into synchronized and non-synchronized. Consequently, values of weights directly related to information about differences between both sets.
The probability density of noise was different in frequencies of measured spectra. The frequencies with the high weights had higher power of spectral lines than the frequencies with the small weight. This fact is described by a mean cumulative distribution function (CDF) of normalized noise of 10 frequencies with the highest and the lowest weights; see Figure 9.

The 90% percentiles ($k_{0.9}$) of the CDF show a significant difference between the discrete frequencies. The frequencies with the lowest percentiles correspond with the frequencies that have highest weights. The five most significant frequencies were found as an intersection of the 25 most important frequencies from the MLNN and CDF analyses. Frequencies of the highest weight and the lowest percentiles were 401, 702, 939, 1528 and 1573 Hz.

6. Conclusion
We used data kindly provided by authors of Jelinek et al. [4]. Process of the increasing electrodynamic activity of the yeasts in the metaphase and anaphase was described by Pohl in electrophoresis experiment [8]. A similar increasing of average power was measured on synchronized cells between 5th and 20th minutes of the measurement. The increasing was caused by a combination of an electrodynamic activity and thermal noise that is depend on an impedance of a cells suspension. A dynamic process of the mitosis changes a distribution and a shape of the cells under the sensor. It may cause changes of the impedance.

The frequency analyses using the data mining method and the CDF of noise shows on 401, 702, 939, 1528 and 1573 Hz. These frequencies were the most significant for classification into two classes.
(synchronized and non-synchronized). Same frequencies had distribution of noise with the greatest occurrence of high power spectral lines. Pelling measured – and Jelínek partially confirmed – the mechanical quasi-coherent oscillations on surface of yeast in acoustic frequency ranges for few seconds using atomic force microscope (AFM) [4, 9]. The mechanical oscillations of electrically polar structures may be one of the sources of the electromagnetic emission that was founded in discrete frequencies.

The results could be distorted by several aspects. The methods assumed independence of the oscillation on a temperature that is not completely fulfilled. The measured system stabilize temperature 28±0.5°C that could cause frequency shift of electrodynamic activity. Very weak signals like thermal noise were measured; so a part of the discovered oscillations may be caused by external noise. However influences of the aspects have not yet been observed.

Acknowledgement

I thank J. Pokorny for valuable discussions and for providing measurement data. This work has been supported by the grants IGA NT11460-4/2010, for which analysing method was primary developing.

Reference

[1] Cifra M, Fields J Z and Farhadi 2011 Electromagnetic cellular interactions Progress in Biophysics and Molecular Biology 105 223

[2] Havelka D, Cifra M, Kučera O, Pokorný J and Vrba J 2011 High-frequency electric field and radiation characteristics of cellular microtubule network Journal of Theoretical Biology 286 31

[3] Kučera O, Cifra M and Pokorný J 2010 Technical aspects of measurement of cellular electromagnetic activity European Biophysics Journal 39 1465

[4] Jelínek F, Cifra M, Pokorný J, Vaniš J, Šimša J, Hašek J and Frýdlová I 2009 Measurement of electrical oscillations and mechanical vibrations of yeast cells membrane around 1 kHz Electromagnetic Biology and Medicine 28 223

[5] Janča R 2010 Measurement of electromagnetic oscillations of cells (in Czech: Měření elektromagnetických oscilací buněk) Faculty of Electrical Engineering, Czech Technical University in Prague

[6] The MathWorks, Inc., © 2010 Neural Network Toolbox. MATLAB documentation

[7] Sakar A and Mammone R J 1993 Growing and pruning neural tree networks IEEE Transactions on Computers 42 291

[8] Pohl H A 1980 Oscillating fields about growing cells Int. J. Quant. Chem. Biol. Symp. 7 411

[9] Pelling A E, Sehati S, Gralla E B, Valentine J S and Gimzewski K 2004 Local Nanomechanical Motion of the Cell Wall of Sacharomyces cerevisiae Science 305 1147