

Markov chain for heart failure diagnostic decision model with irregular Thoracic Fluid Content measurements

Paweł Moszczyński, Andrzej Walczak

Institute of Computer and Information Systems, Military University of Technology, Warsaw, gen. Urbanowicza St.2, Poland

Abstract. Properties of results obtained on Markov processes analysis in medical applications have been illustrated. Calculation was done for measurement of Thoracic Fluid Content by means of impedance cardiograph for chosen medical patients population. Main medical aspect of Markov chain calculation has been underlined to exhibit some constraints in decision making with such math models. Markov chain analysis was done as optimization task. Effective algorithm for irregular data in continuous time has been proposed.

1 Introduction

Each tissue of the human body is characterized by electrical resistance depending on tissue density, temperature, and molecular composition. Blood conducts electricity better (R blood plasma is around 65 Ω/cm, whole blood is around 130 Ω/cm) than any other thoracic tissues such as muscles, bones, lungs and adipose tissue (R is in the range 200-5000 Ω/cm) [1]. Therefore, the impedance measurement (high frequency, low magnitude current is transmitted through the chest) allows the assessment of many important hemodynamic parameters related to the blood flow in a given body segment and its cyclical changes.

In cardiography the placement of four dual disposable sensors on the neck and chest are used to detect the impedance changes caused by a high-frequency, low magnitude current flowing through the thorax [2, 3]. Using data from the impedance cardiograph (ICG) device, we can measure or calculate heart parameters such as: HR (Heart Rate), TFC (Thoracic Fluid Content), SBP (Systolic Blood Pressure), DBP (Diastolic Blood Pressure), Delta TFC (difference between two next values of TFC) and others. The Markov models which are presented in the publication describe evolution of the patient's state determined from the TFC [1/kΩ] parameter (see figure 1). However, the method of proceeding can be applied to all ICG parameters used in medical diagnosis.

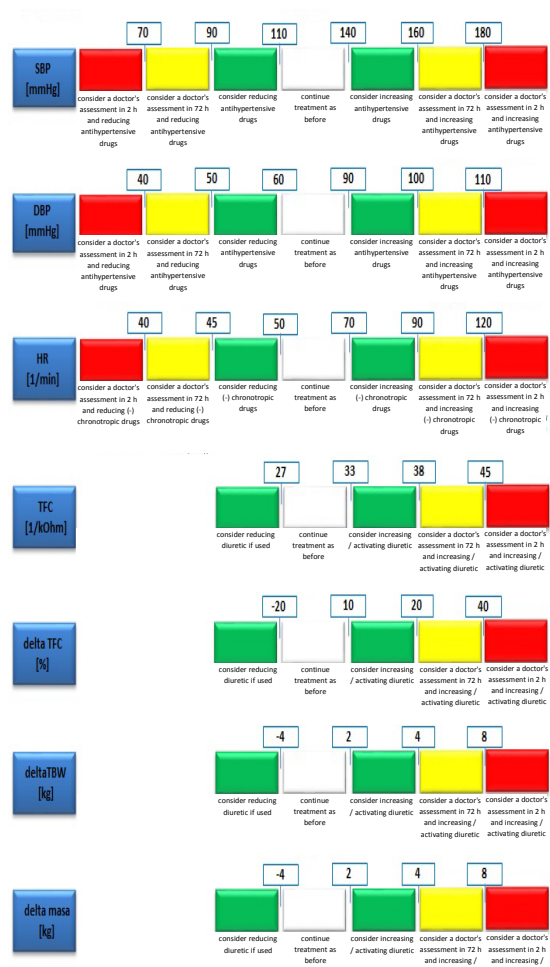


Fig. 1. Examples of measured data for cardiology. Each section inside measured value (green, white, yellow, red flag) means in fact separated condition of the observed patient. [2]

Basic form of Markov chain is applied in our model it means just chain with Markov attribute.

All measurements and analysis are part of project exploiting impedance cardiogram (ICG) for heart failure diagnostic [4]. Observed patients population is not so large yet but common feature of observation is that time distance for each measurement inside patient population are different, and irregular. It is the reason we propose couple of Markov chain model to exhibit their differences.

It seems to be quite important question if Markov processes describe properly prediction for medical patient state evolution. First of all medical knowledge is imposed on our calculation with any connection to Markov processes. Second - if we have a look at figure 1 one can see state space (or phase space of Markov process if you will) divided in accordance with actual medical knowledge of TFC observation. So space dimension equal to 5 is imposed to our calculation. On the other hand we have population of different patients with individual properties not obligatory equivalent to 5 dimensional space of state.

Simultaneously it is also well known, that result of calculation based on Markov chains strongly depends on number of states used inside chain construction. Used geometry of chain as well as its dimension, and jumps inside states taken into account influences on final values of calculated probabilities.

Nevertheless in 1983, Beck and Pauker described the use of Markov models for determining prognosis in medical applications [3]. Since that introduction, Markov models have been applied with increasing frequency in published decision analyses [5-10] to mention only a few which started big number of similar analysis. Most of publications were joined with decision tree evolving in examined illnesses. So space of decision was natural consequence of contemporary medical procedures.

There is really big amount of papers on Markov processes so we decides point out only basic connections in models for medical decisions.

In TFC measurement one can assume that TFC range divided into 5 parts is also result of medical procedure. Unfortunately set of ICG measurement made on patient population is rather random, and

irregular. It seems to us worth to analyze such situation in possible suitable way with Markov chain tool.

Structure of document consists of five parts including introduction. Part 2 describes basic rules assumed for patient data model. Parts 3, 4, 5 contain steps in data analysis with Markov model for the same intervals between measurement, for different interval, and finally for irregular intervals with imposed probability distribution as one of the constraints in optimization.

2 Application of Markov chains in the description of the evolution of the patient's state

We assumed that the Markov chain is a sequence of random variables $X_{m,1}, X_{m,2}, X_{m,3} \dots$ describing the state of the m -th patient in 1,2,3 and subsequent times. We assumed also that the Markov chain for each patient is homogeneous.

In addition, due to the small number of data for individual patients, they were combined into groups. For each group, the matrixes of transitions between states for patients in a given group were assumed identical. As a result, the number of data used to determine the parameters of the chain describing the evolution of the patient's condition from a given group has been increased.

In the first model, we assume that further tests for patients in a given group are carried out at regular intervals. So variable $X_{m,n}$ describe the state of the m -th patient in the n -th test.

Second model describes how patient's condition changes within one day. In this model, the random variable $X_{m,n}$ describes the state of the m -th patient on the n -th day. In this approach random variable $X_{m,n+k}$ describes that test for the m -th patient took place on the n -th day and the next will take place in k days.

As the third, we present a model that is an extension of the model describing how the patient's condition changes within one day. In this approach, a constraint was added to the optimization problem that the distribution of probabilities after the endless number of days tends to a particular, chosen probability distribution. This approach allows to increase the number of states in which the patient is staying, and therefore to increase the sensitivity of the prognosis.

To obtain value of probability we create one dimensional Markov chain, and for each mentioned model we are looking for optimal solution that the sum of logarithms of probabilities occurrence of the observed data achieves maximum value.

3 Model for the same intervals between successive events

The model presented in this chapter is applicable if further tests for patients in a given group are carried out with regular intervals.

Let us assume that average value of the TFC parameter for a single patient changed as in Figure 2.

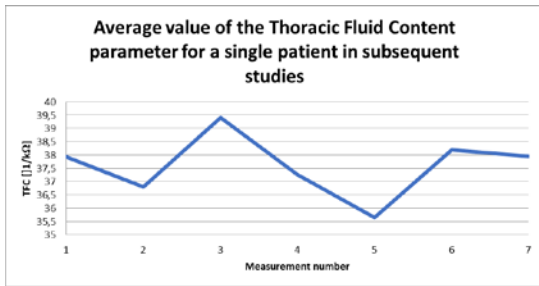


Fig. 2. Sample measurements of fluid in the patient's chest in subsequent tests.

Then, in accordance with Figure 1, we take into account the following transitions between the states:

$$(S_4 ; S_3), (S_3 ; S_4), (S_4 ; S_3), (S_3 ; S_3), (S_3 ; S_4), (S_4 ; S_3),$$

The transition matrix for the Markov chain describing the evolution of the patient is a square matrix of the order S (number of states in Markov chain) and depends on the number of states for a given highlighted parameter. In the case of TFC S=5 (see figure 1). When determining individual probabilities in the matrix, we use the fact that the chain is homogeneous. In this case, the probability of transition from state i to state j is equal to the quotient l_{ij} the sum of transitions from state i to state j and the sum of transitions from state i to any state. For TFC, we get the following formula:

$$p_{ij} = \frac{l_{ij}}{\sum_{k=1}^S l_{ik}}. \quad (1)$$

Because we assumed that for the remaining patients in the analyzed group, the Markov chain contain an identical matrix of

transitions, while adding up, we also consider transitions between states for all patients in a given group.

4 Model for various intervals between successive events

If the tests are performed every certain number of days, the model described in the previous chapter cannot be used. By monitoring the health of patients with heart failure, the intervals between subsequent tests range from one to several dozen days. In this chapter, we present model describing how the patient's condition changes within one day. For the illustration of model building, let us assume that we only consider periods of time not longer than 14 days ($k = 14$).

Let l_{ijk} denote the sum of transitions from the i-th state to the j-th state in k days. Analogously to the model in the previous chapter, we sum up the transitions measured in patients from the analyzed group.

To simplify calculations, we assume that in one day the state may change by one or not at all. This assumption results from the fact that the appropriate l_{ij1} coefficients are zero in the experimental data we analyze. The matrix describing the Markov process is as follows:

$$P = \begin{bmatrix} p_{11} & p_{12} & 0 & 0 & 0 \\ p_{21} & p_{22} & p_{23} & 0 & 0 \\ 0 & p_{32} & p_{33} & p_{34} & 0 \\ 0 & 0 & p_{43} & p_{44} & p_{45} \\ 0 & 0 & 0 & p_{54} & p_{55} \end{bmatrix} \quad (2)$$

The probability of transition from the state i to the state j in k days p_{ijk} we obtain from the Chapman-Kolmogorov equation and is:

$$p_{ijk} = s(i)P^k s^T(j), \text{ where } i, j, k \in \text{and} \quad (3)$$

$$1 \leq i, j \leq 5.$$

In formula (3) $s(i)$ is a vector filled with zeros and one in i-th place and P^k is the transition matrix multiplied by k-fold. Therefore, the objective function and restrictions are as follows:

$$\max_P \sum_{k=1}^K \sum_{i=1}^S \sum_{j=1}^S l_{ijk} \log(s(i)P^k s^T(j)) \quad (4)$$

Where parameters of the matrix P must follow constraints:

$$\begin{aligned}
 p_{11} + p_{12} &= 1 \\
 p_{21} + p_{22} + p_{23} &= 1 \\
 p_{32} + p_{33} + p_{34} &= 1 \\
 p_{43} + p_{44} + p_{45} &= 1 \\
 p_{54} + p_{55} &= 1 \\
 \varepsilon \leq p_{11}, p_{12}, p_{54}, p_{55} &\leq 1 - \varepsilon \\
 \varepsilon \leq p_{w,w-1}, p_{ww}, p_{w,w+1} &\leq 1 - \varepsilon
 \end{aligned} \tag{5}$$

*the remaining factors
 of the matrix P are equal to zero*
 $w \in \mathbb{N}$, gdzie $2 \leq w \leq 4$

In the above model, the ε parameter is a small number that has been introduced to make $p_{ijk} \neq 0$ for all $l_{ijk} \neq 0$. Due to the fact that we are dealing with a smooth and strongly non-linear optimization problem (4-5), a generalized reduced gradient method [11,12,13, and 14] was chosen for solving it.

Using the matrix P determined by solving the optimization problem (4,5), from the formula (3) we can determine the probability of finding each state after k days. The dependence of the probability of a jump from the second state (the TFC value is in the range of 27-33 [1 / kΩ]) to individual states on the number of days is shown in Figure 3. The results based on 120 tests performed by 32 patients.

number of days	probability of transition to state 1	probability of transition to state 2	probability of transition to state 3	probability of transition to state 4	probability of transition to state 5
1	0,123203	0,787132	0,089665	0,000000	0,000000
2	0,208754	0,643093	0,137911	0,010242	0,000000
3	0,268624	0,544156	0,161753	0,025467	0,000000
4	0,310753	0,475046	0,171571	0,042629	0,000000
5	0,340460	0,425879	0,173632	0,060028	0,000001
6	0,361354	0,390211	0,171669	0,076765	0,000001
7	0,375916	0,363809	0,167859	0,092414	0,000002
8	0,385875	0,343863	0,163438	0,106820	0,000003
9	0,392453	0,328491	0,159073	0,119979	0,000004
10	0,396527	0,316413	0,155095	0,131960	0,000005
11	0,398735	0,306747	0,151643	0,142869	0,000007
12	0,399547	0,298878	0,148746	0,152821	0,000008
13	0,399314	0,292369	0,146379	0,161929	0,000010
14	0,398301	0,286904	0,144487	0,170296	0,000011

Fig. 3. Probability of jump into particular states of the patient whose TFC was measured from the interval 27-33[1/kΩ] after the number of days shown in left column.

5 Model with stationary distribution determined from the assumed distribution

In this approach to the optimization problem (4, 5), the condition was added that the distribution of probabilities at the endless number of days (stationary distribution) tends to a particular

distribution whose parameters are determined experimentally.

It allows to increase the number of states in which the patient is staying, and therefore to increase the sensitivity of the prognosis. Due to the smaller range of the measured parameter attributable to a single state, we assume that after one day the state can change by two, one or remains unchanged. For example, the matrix of transitions for S=9 is as follows:

$$P = \begin{pmatrix}
 p_{11} & p_{12} & p_{13} & 0 & 0 & 0 & 0 & 0 & 0 \\
 p_{21} & p_{22} & p_{23} & p_{24} & p_{25} & 0 & 0 & 0 & 0 \\
 p_{31} & p_{32} & p_{33} & p_{34} & p_{35} & 0 & 0 & 0 & 0 \\
 0 & p_{42} & p_{43} & p_{44} & p_{45} & p_{46} & 0 & 0 & 0 \\
 0 & 0 & p_{53} & p_{54} & p_{55} & p_{56} & p_{57} & 0 & 0 \\
 0 & 0 & 0 & p_{64} & p_{65} & p_{66} & p_{67} & p_{68} & 0 \\
 0 & 0 & 0 & 0 & p_{75} & p_{76} & p_{77} & p_{78} & p_{79} \\
 0 & 0 & 0 & 0 & 0 & p_{87} & p_{87} & p_{88} & p_{89} \\
 0 & 0 & 0 & 0 & 0 & 0 & p_{97} & p_{98} & p_{99}
 \end{pmatrix} \tag{6}$$

In addition, the presented method allows solving the problem of the lack of results in critical states. For TFC, it is assumed that a critical condition is one in which the TFC value is over 45 [1/kΩ]. According to the medical recommendations patient in such a state should urgently contact hospital within 2 hours to a significant change in treatment. However, available results show that patients who have been tested are very rarely in critical condition. On the other hand, the annual mortality rate of people with heart failure is respectively in NYHA classes: class I - to 10%, class II - 10–20%, class III - 20–40%, and class IV - mortality 40–60%[2]. Different mortality in individual groups is therefore included in stationary distributions established for these groups.

The objective function and restrictions are as follows:

$$\max_P \sum_{k=1}^K \sum_{i=1}^S \sum_{j=1}^S l_{ijk} \log(s(i)P^k s^T(j)). \tag{7}$$

Where parameters of the matrix P for S=9 must follow constraints:

$$\begin{aligned}
 p_{11} + p_{12} + p_{13} &= 1 \\
 p_{21} + p_{22} + p_{23} + p_{24} &= 1 \\
 p_{31} + p_{32} + p_{33} + p_{34} + p_{35} &= 1 \\
 p_{42} + p_{43} + p_{44} + p_{45} + p_{46} &= 1 \\
 p_{53} + p_{54} + p_{55} + p_{56} + p_{57} &= 1 \\
 p_{64} + p_{65} + p_{66} + p_{67} + p_{68} &= 1 \\
 p_{75} + p_{76} + p_{77} + p_{78} + p_{79} &= 1 \\
 p_{86} + p_{87} + p_{88} + p_{89} &= 1 \\
 p_{97} + p_{98} + p_{99} &= 1
 \end{aligned} \tag{8}$$

$$r(\beta)P=r(\beta)$$

$$\varepsilon \leq p_{11}, p_{12}, p_{13}, p_{21}, p_{22}, p_{23}, p_{24} \leq 1 - \varepsilon$$

$$\varepsilon \leq p_{86}, p_{87}, p_{88}, p_{89}, p_{97}, p_{98}, p_{99} \leq 1 - \varepsilon$$

$$\varepsilon \leq p_{w,w-2}, p_{w,w-1}, p_{ww}, p_{w,w+1}, p_{w,w+2} \leq 1 - \varepsilon$$

the remaining factors of the matrix P are equal to zero
 $\alpha - \Delta \leq \beta \leq \alpha + \Delta$
 $w \in \mathbb{N}$, where $3 \leq w \leq 7$

In the above model, a vector r is stationary distribution for the searched Markov chain. Probabilities occurring in it are determined based on the certain assumed distribution. At the same time, we assume that the parameter of this distribution β oscillated in the range determined on the basis of clinical results $(\alpha - \Delta, \alpha + \Delta)$.

Figures 4, and 5 illustrates the probability of getting into individual medical states (see figure 1) if during the examination TFC was within the range 27-29 [1/kΩ] (figure 4) or 31-33[1/kΩ] (figure 5). The results were obtained on the basis of the P matrix which is the solution of the formulae (7-8), in which number of state in Markov chain is equal 15 and the factors l_{ijk} , α were obtained on the basis of 120 tests performed on 32 patients.

number of days	probability of transition to state 1	probability of transition to state 2	probability of transition to state 3	probability of transition to state 4	probability of transition to state 5
1	0,293922	0,706078	0,000000	0,000000	0,000000
2	0,369752	0,598842	0,031406	0,000000	0,000000
3	0,382669	0,551033	0,065786	0,000512	0,000000
4	0,385602	0,525557	0,087228	0,001613	0,000000
5	0,386955	0,509622	0,100031	0,003375	0,000018
6	0,387619	0,497756	0,108918	0,005637	0,000069
7	0,387960	0,487887	0,115806	0,008186	0,000160
8	0,388131	0,479343	0,121374	0,010861	0,000291
9	0,388123	0,471882	0,125973	0,013564	0,000457
10	0,387899	0,465364	0,129850	0,016234	0,000653
11	0,387443	0,459673	0,133182	0,018830	0,000872
12	0,386765	0,454702	0,136094	0,021331	0,001108
13	0,385888	0,450356	0,138680	0,023721	0,001355
14	0,384840	0,446552	0,141006	0,025994	0,001607

Fig. 4. Probability of jump into particular states of the patient whose TFC was measured from the interval 27-29[1/kΩ] after the number of days in left column.

number of days	probability of transition to state 1	probability of transition to state 2	probability of transition to state 3	probability of transition to state 4	probability of transition to state 5
1	0,000000	0,917604	0,082396	0,000000	0,000000
2	0,037825	0,765307	0,192219	0,004649	0,000000
3	0,101919	0,671393	0,217508	0,009180	0,000000
4	0,152936	0,617689	0,214768	0,014513	0,000096
5	0,188528	0,581628	0,209360	0,020165	0,000319
6	0,214984	0,553963	0,204903	0,025502	0,000647
7	0,236034	0,531612	0,201035	0,030272	0,001048
8	0,253268	0,513210	0,197583	0,034450	0,001489
9	0,267526	0,497892	0,194549	0,038085	0,001947
10	0,279394	0,485045	0,191918	0,041239	0,002403
11	0,289319	0,474214	0,189649	0,043974	0,002844
12	0,297647	0,465050	0,187696	0,046345	0,003262
13	0,304651	0,457274	0,186019	0,048405	0,003651
14	0,310550	0,450662	0,184583	0,050195	0,004010

Fig. 5. Probability of jump into particular states of the patient whose TFC was measured from the interval 31-33[1/kΩ] from the number of days.

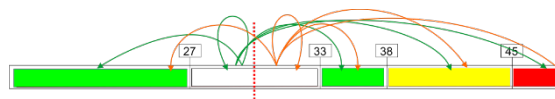


Fig. 6. Scheme of transitions. Figure 4 represents probability of transition from left side of state white (range: 27-29 [1/kΩ]) to all states (green arrows). Figure 5: probability of transitions from right side of state white (range: 31-33 [1/kΩ]) (orange arrows).

6 Resume

Results illustrate that Markov processes for irregular TFC measurement can indicate different values in dependency of time marker model. All results have been created as goal function maximization during classical optimization task, and resolved with the same method for properly constructed constraints. So in fact only constraints decide about final probabilities. We get time series of 14 days duration, and probabilities are significantly different in dependency on boundaries with the same goal function. Results are for the same patient population, and for the same TFC measurement. Some similarities can be observed in results especially, that state 2 is preferred in all calculations, and some specific asymmetry in preferred states is observed. We do suppose that such kind of stability in calculations proofs their correctness to some extent.

It seems to us that Markov chains must be applied in medical decision support carefully, and after deep clinical studies. Especially sensitivity of such calculation should be examined during clinical tests.

Acknowledgements

This study was supported by National Centre of Research and Development, STRATEGMED3/305274/8/NCBR/2017.

References

1. P. Krzesiński, G. Gielerak, J. Kowal Impedance cardiography - a modern tool for the monitoring of cardiovascular diseases, *Kardiol Pol* 2009; 67: 65-71
2. HH. Woltjer, HJ. Bogaard, PM. de Vries The technique of impedance cardiography *Eur Heart J* 1997; 18: 1396-403
3. JM. Van De Water, TW. Miller, et al. Impedance cardiography: the next vital sign technology? *Chest* 2003; 123: 2028-33
4. <https://amulet.wim.mil.pl>
5. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making*. 1983;3:419-58.
6. Eckman MH, Beshansky JR, Durand-Zaleski I, Levine HJ, Pauker SG. Anticoagulation for noncardiac procedures in patients with prosthetic heart valves. Does low risk mean high cost? *JAMA*, 1990;263:1513-21.
7. Wong JB, Sonnenberg FA, Salem D, Pauker SG. Myocardial revascularization for chronic stable angina: an analysis of the role of percutaneous transluminal coronary angioplasty based on data available in 1989. *Ann Intern Med*. 1990;113:852-71.
8. Hillner BE, Smith TJ, Desch CE. Efficacy and cost-effectiveness of autologous bone marrow transplantation in metastatic breast cancer. Estimates using decision analysis while awaiting clinical trial results. *JAMA*. 1992;267:2055-61.
9. Birkmeyer JD, Marrin CA, O'Connor G-T. Should patients with Bjork-Shiley valves undergo prophylactic replacement? *Lancet*. 1992;340:520-3.
10. Sonnenberg F., Beck R. *Med Decis Making* 1993;13:322-338), <http://sagepub.mdm.com> at National Institutes of Health Library on January 21, 2009
11. L. S. Lasdon, A. D. Waren, A. Jain, and M. Ratner. 1978. Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming. *ACM Trans. Math. Softw.* 4, 1 (1978), 34-50.
12. J. Abadie and J. Carpentier, "Generalization of the Wolfe Reduced Gradient Method to the

- Case of Nonlinear Constraints," in *Optimization* R. Fletcher, Ed., Academic Press, 1969, 37-47.
13. J. Abadie, "Application of the GRG Algorithm to Optimal Control Problems," in *Nonlinear and Integer Programming* J. Abadie, Ed., North Holland Publishing Co., 1972, 191-211.
 14. L. S. Lasdon, R. Fox and M. Ratner, "Nonlinear Optimization Using the Generalized Reduced Gradient Method," *Tech. Memo. No. 325*,