

# Cloud technologies in the problems of mathematical analysis of cardiological information

E Zimina<sup>1,2</sup>, M Novopashin<sup>2</sup> and A Shmid<sup>1,2</sup>

<sup>1</sup>National Research University Higher School of Economics, Myasnitskaya str. 20, Moscow, Russia, 101000

<sup>2</sup>EC-Leasing, Varshavskoe shosse 125/1, Moscow, Russia, 117587

**Abstract.** The article includes the observation of the cloud services and technologies usage. The article contains a review of mathematical analysis of cardiac information using cloud technology, which produces storage, analysis and forecasting on the basis of owned data. In addition, the authors consider the possibility of integrating cloud technologies with external systems. The massive use of mobile devices for the removal of the electrocardiogram (ECG) leads to a quantitative increase of the patients number available for ECG investigation. Thus, there are new opportunities to research the oscillatory processes of long-term dynamics of the individual state of the cardiovascular system (CVS) of any patient. The article demonstrates new opportunities the long-term continuous monitoring of the patients CVS, which allows identifying regularities of the dynamics of the CVS. Also this article comprises the observation of the existence of an adequate model of CVS as a distributed nonlinear self-oscillating system of the class model returns the Fermi-Pasta-Ulam (FPU).

## 1. Introduction

According to the forecast of the McKinsey Institute for the 2025 year in the field of digital economy the most significant areas of technologies application are Mobile Internet, Automation of knowledge work, Internet of Things and Cloud. The total market volume will be around 30 trillion dollars. For medical use of these technologies the estimated volume is approximately 30%. For comparison, the share of oil and gas sector of the world economy is estimated at only \$ 1.5 trillion.

Also, according to the Google Academy (Google Scholar) there is a sharp increase in the number of publications with reference to Big Data among all publications on research in the field of health [1].

The aim of the project is to create a centralized system of cardiological information processing of the population outside the medical institutions, providing personalized monitoring, analysis and forecasting of the development of the patient's condition in real time before his first appearance in the hospital.

This project is one possible response to the global health challenge: on the one hand, the growing global contradiction between the existing medical capabilities of the world health system and, on the other hand, the shortcomings in the concept and technologies of its organization, leading to a rapid increase in the lack of all types of resources in its application [2].

The primary component of the system of cardiological information processing in this project is the cloud, which combines in a single automated system and mobile devices, a group of trained computers and medical personnel who make decisions on scenarios of interaction with patients.

A system of processing of cardiac information is being developed, which provides personalized monitoring, analysis and forecasting of the patient's condition in real time.

## 2. The concept of FPU auto-return

We introduce the notion of FPU auto-return at the mathematical level. The phenomenon of canonical Fermi-Pasta-Ulam (FPU) auto-return in passive systems was the first described as a result of numerical study of differential solutions – difference equations describing a chain of nonlinearly coupled vibrators [3].

In these model chains the phenomenon of attenuation has not been submitted. To describe real dynamic processes, such as cardiac activity, a more appropriate model is needed that can describe returns in autonomously functioning or self-oscillating systems.

A significant contribution in this direction was the work of American researchers Zabuski and Kruskal [4], who proposed to describe the FPU auto-return in the framework of the Korteweg de Vries equation (KdV) with periodically changing boundary and initial conditions.

Using the results of this work it is possible to simplify the solution of the KdV equations in the form knoid waves, replacing them with the solutions of the equation of Van der Pol, close to harmonic, periodic and low frequency changes in boundary and initial conditions on the relaxation solution of the equation of Van der Pol [5].

Meanwhile, if there is the theorem application of the replacing wave links with lagging ones possibility, it is possible to present a mathematical model of the electrical activity of the heart (ECG) in the form of FPU auto-return, described in the framework of Van der Pol equations with delay in the following form [6]:

$$\frac{d^2M_1}{dt^2} - a_1(1 - Y_1) \frac{dM_1}{dt} + \omega^2(1 + a_1M_2)M_1 = c_1 \frac{d^2M_2}{dt^2} + d_1F_1 + d_2F_2 \quad (1)$$

$$b_1Y_1 + T_1 \frac{dY_1}{dt} = M_1^2 \quad (2)$$

$$\frac{d^2M_2}{dt^2} - a_2(1 - Y_2) \frac{dM_2}{dt} + \omega_2^2(1 + a_2M_1)M_2 = c_2 \frac{d^2M_1}{dt^2} + d_1F_1 + d_2F_2 \quad (3)$$

$$b_2Y_2 + T_2 \frac{dY_2}{dt} = M_2^2 \quad (4)$$

where  $M_1$  - the value proportional to a dynamic electrical potential of the myocardium,  $M_2$  - the value proportional to the dynamic electric potential is a small locality of the myocytes at the surface of the myocardium,  $Y_1$  - the value proportional to the delay in the propagation of electrical impulse in the myocardium,  $Y_2$  - the value proportional to the delay in the propagation of electrical impulse in the locality of myocytes,  $b_1$  - the value proportional to the surface areas of the myocardium,  $b_2$  - the value proportional to the surface area of the locality of myocytes,  $T_1$  - the value proportional to the time of myocardial contraction,  $T_2$  - the value proportional to the period of oscillation in the locality of myocytes,  $F_1$  - the function of the resonant effect of the external environment on the heart rate (about 1 Hz),  $F_2$  - the function of the resonant effect of the external environment on the heart at a high frequency of about 20 Hz,  $c_1, c_2, d_1, d_2$  - constants.

The properties of ECG spectrum of healthy people and patients with ischemic heart disease found in the course of the study will allow to formulate the problem of mathematical modeling of the heart electrical dynamics with the description of the phenomena observed in the experiment [7] so that the heart can be represented as an open dynamic structure with many self-oscillating degrees of freedom, as shown in formulas 1-4.

The approaches proposed by Norbert Wiener, Ginsburg and Landau have been successful in modelling the dynamic states of cardiac electrical activity such as normal functioning, ischemic disease and myocardial infarction.

## 3. The Cloud Technologies

The cloud allows us to receive, store and process the electrocardiogram, photoplethysmogram (pulse wave), echocardiograms, seismocardiogram for further calculations. A first lead is required for in-depth analysis, but any additional leads only clarify further analysis.

The cloud is replenished by means of application programming interface (API) and adapters of two types: intended for self-unloading of data and intended for reading ECG formats received from specialized devices.

The cloud is designed for storage and processing using Big Data technologies information obtained from specialized agencies and from the mobile user's device.

The results of the analysis are also available via the API and special adapters to integrate with external systems.

Calculations of indicators are made on the basis of frequency and energy characteristics of the received signals. There is a selection of typical and transient frequency states for each patient individually, as well as for target samples of signals, divided by age, sex and the presence of pathologies in the patient.

The analysis determines the proximity of the calculated characteristics of the patient to one of the certain frequency states according to the data of one examination. Tracks the progress indicators of the patient status between the model states of the patient and the target samples to determine trends in the development of pathologies, effects of medications, post-operative rehabilitation.

The biological age of the patient's cardiovascular system is determined as a separate type of trajectory. There is an opportunity to detect acute the patient's condition that requires professional help.

By lengthening the trajectory of the calculated characteristics of the patient's condition between the typical patient state and the target samples, prediction of the future states of the patient for the next few days is constructed.

When a new cardiogram is entered into the database, two types of treatment are initiated: cardiogram treatment for the current patient and processing of an updated set of all cardiograms from the database.

## 4. Work with cardiograms

### 4.1. *Cardiograms handling*

The investigation of cardiograms is necessary to obtain calculations on the basis of which the analysis of the patient's condition and prognosis of his condition will be made. Processing of a new incoming cardiogram is performed in the sequential execution of several types of tasks:

- preparation of cardiogram for further calculations;
- calculation of the Fourier spectrum and its characteristics (harmonic and energy) for the unfiltered signal, signal with the corrected contour and signal with the corrected contour and free from defects;
- calculation of spectral shapes for the signal with the corrected contour and free from defects;
- recalculation of cardiogram clusters in the entire series of the patient cardiogram with different clustering parameters;
- calculation of the patient trend indicators (biological age).

The updated set of all cardiograms is processed according to the schedule and for different target groups by age, sex, pathology information for various clustering parameters.

Age groups are as follows: all ages and 5 separate age groups from young to old. Gender groups are as follows: men and women.

The following groups are distinguished by pathologies:

- spaced sampling cardiogram diagnosis of ischemia (confirmed and not confirmed);
- marked sample of cardiogram diagnosis of tuberculosis (confirmed and not confirmed);
- marked selection of cardiograms diagnosed diabetes (confirmed and not confirmed).

If there is not new ECG since the last clustering, the launch of the new clustering is not performed.

In the processing of individual prior ECG to performing any calculations and research, it is necessary to correct the contour, i.e. to remove the drift contours, and clear it from the strong defects, because it spoils the frequency response of the required ECG is needed for further analysis [8].

Drift contours removed with the help of the two median method, from original signal elementwise subtracted calculated by the two median contour.

#### 4.2. *Cardiogram preparation for calculation of indicators*

Before processing a separate cardiogram, it is necessary to remove the drift of the contour, and clean it from strong defects.

Drift contours removed two median method, from original signal elementwise subtracted calculated by the two median contour. Defect filter is used. RR intervals are cut out from the signal, in which the signal amplitude spread is much higher than the value of the median amplitude spread over a sliding window [9]. Parts of the signal are glued together at the vertices of R-peaks, between which the RR-interval was cut. The edge parts before the first detected R-peak and after the last detected R-peak are also discarded so that all filtered signals start at R-peak and end at R-peak.

The percentage of the dropped part of the signal in relation to the entire length of the signal is stored separately.

### 5. Calculation of Fourier spectrum and its characteristics

#### 5.1. *Methods of research*

Patients with permanent atrial fibrillation were excluded from the study, as well as the following patients were excluded: with complete atrioventricular blockade, complete blockade of the left leg of the bundle branch block, the technical causes of false negative or false positive detection in the identification of myocardial ischemia (recommendations for Holter), life-threatening arrhythmia, acute cerebrovascular accident, tumors, severe anemia, thyroid disease in the stage decompensation, lack of informed consent of the patient, failure to conduct stress tests and with questionable results of stress tests and Holter monitoring (HM) ECG [10].

The diagnosis was based on clinical, anamnetic, instrumental methods of research. Verification of the diagnosis was carried out using samples with physical activity (veloergometry or stress echocardiography). All patients underwent HM ECG.

#### 5.2. *Calculation of Fourier spectrum*

For an unfiltered signal, a signal with a corrected isolation, and a signal with a corrected isolation and filtered defects, a fast Fourier transform algorithm is used. The amplitude-frequency spectrum is calculated from the complex Fourier transform coefficients. On the spectrum, harmonics are defined – characteristic “humps” in the region of certain frequencies. The amplitude and frequency of each harmonic are determined inside the window calculated based on the base frequency. The base frequency is the heart rate displayed in Hertz, i.e. beats per second. It is calculated as follows:

$f_{\text{base}} = 1/\text{median}(\{\text{RRs}\})$ , where  $\{\text{RRs}\}$  is the set of RR interval lengths.

For each harmonic, the frequency  $f_i$ , amplitude  $A_i$  and energy index – the spectral density  $E_i$  are calculated.

### 6. Clustering of cardiograms

#### 6.1. *Clustering of cardiogram samples*

Clustering is performed for a variety of cardiograms, isolated by certain characteristics. For the tasks clustering is performed for all individual patient's cardiograms with different clustering parameters according to the selected forms of the spectrum.

The distance value from the shape of the spectrum to the cluster is considered as the distance from the shape of the spectrum to the cluster middle line, which is always stored as an intermediate result. Merging of clusters is accompanied by a recalculation of the values of the average lines in accordance with the number of cardiogram in each of the merged clusters. Also clustering can be carried out by other methods: on the basis of the form obtained by another algorithm, based on the average cardiocycle and on the basis of amplitudes or spectral densities of harmonics.

#### 6.2. *The calculation of indicators of trends in the patient*

To determine the trend of changes in indicators of the patient's condition between the two states, the most popular clusters are the largest number of cardiograms.

At this stage, a cluster of patients cardiograms of the younger age group and a cluster of patients cardiograms of the older age group, regardless of gender, are distinguished. Next, we will call it the “young” cluster and the “old” cluster, respectively.

The trend index  $t_{(ECG-M)}$  is the ratio of the distance from the current form of the patient's spectrum to the middle line of the “young” cluster to the distance to the middle line of the “old” cluster.

For each trend indicator, the confidence factor  $w_{(ECG-M)}$  calculated from the triangle rule is considered. The most indicative are the patient's conditions located close to the line between the “young” cluster and the “old” cluster.

The trust factor is set to 0 to 1. The higher it is, the more significant the indicator is for the research.

Based on a set of pairs  $\{t_{(ECG-M)}, w_{(ECG-M)}\}$ , the trend indicators for the last day, last 4 days, week (last 7 days), month (last 30 days), and all time of patient observation are determined. They are calculated by the formula of mathematical expectation. On request, the trend indicator can be calculated for an arbitrary period of time.

6 values are calculated for each cardiogram, which are available for external systems.

### 6.3. Building ECG clusters principles

Cluster refers to a set of one or more ECG and associated with this set, the arithmetic mean of the spectrum, calculated automatically as a result of the ECG ordered classification process.

Comparison of ECG Fourier spectrum was implemented on the basis of the FPU auto-return phenomenon with the visual perception of the spectrum. For this purpose, the concept of spectrum shape was introduced. Spectrum shape-smoothed curve based on the median of the upper 30% of the values in the window 2 Hz, taken in the range from 0 to 40 Hz. Window width 2 Hz is taken specially to eliminate the influence of pronounced harmonics for ECG with a heart rate less than 120 beats/min. Form of the spectrum is used further in the algorithm for determining regions of ischemia and clustering of ECG in the form of the spectrum.

Combining similar self-oscillating processes of the heart into clusters on the basis of belonging to the diagnosis as a result give a finite number of subsets. However, it follows from the FPU model that there may be such variants of behavior of the heart that they will not “fall” into any of the clusters, which correspond to the known diagnoses. Such “outliers” reflect the process of the heart model transition from one state (cluster) to another [11].

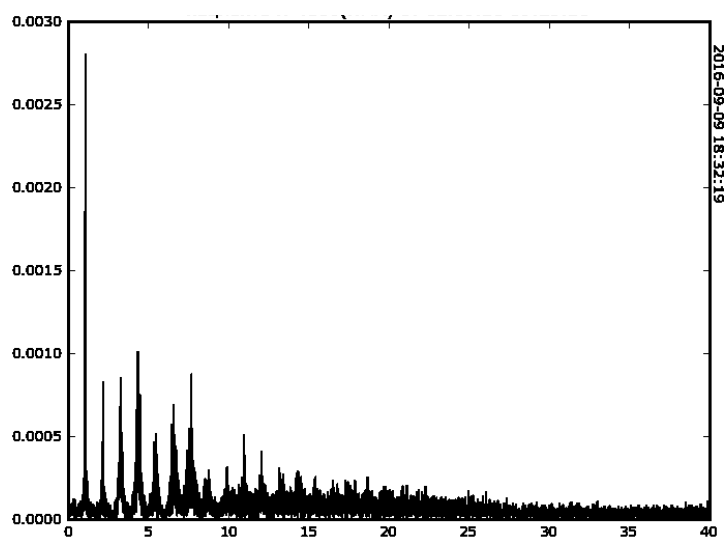
For clustering, the form is approximated in steps of 0.1 Hz to 400 points and normalized to the arithmetic mean for comparison universality.

ECG clustering in the form of spectrum is performed by the method of extraction of connected components. The distance between the forms of the compared ECG spectrum is taken Euclidean. In the process of clustering the original fully connected graph, in case when it links of greater length than the assigned value are dropped, is scattered into a group of subgraphs, which form the desired clusters as a result.

In order to simplify the algorithm clusters are formed automatically based on ECG classification as they are received (the so-called online method). In this case, operations are meant to supplement the cluster with a new form of ECG, it creates a new cluster and merges clusters based on the distance between them.

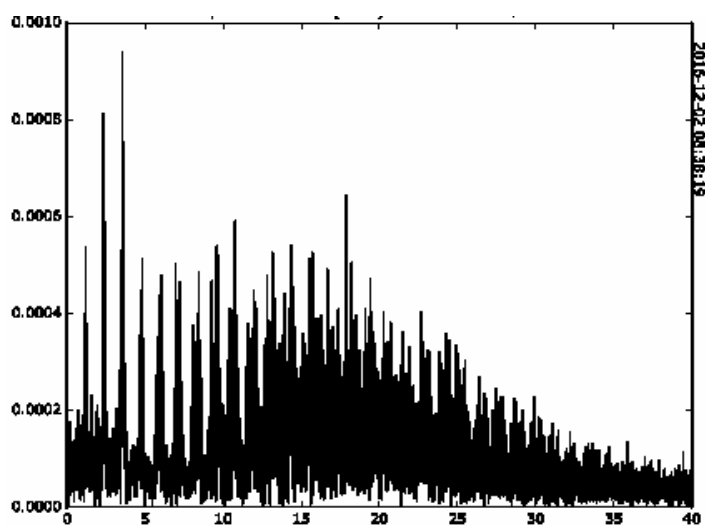
It is practically obtained that there is always a cluster with a predominant number of ECG trapped in it. Presumably, the ECG of this cluster characterize the normal state of the patient. ECG, trapped in the other clusters, characterize to some extent the abnormal condition of the patient, caused by both external stress and pathologies. Often there are clusters consisting of a single ECG, their appearance gives an indication of the intermittency of complex FPU auto-return patient (when between occurrences of a certain state there is a change of more states than from repeated).

For such reference points several clusters of healthy people, patients with ischemia were selected. In the Figure 1 there is a typical Fourier spectrum of the ECG of a healthy person. In the Figure 2 the typical Fourier spectrum of the ECG of a patient with a developed ischemic process is given.



**Figure 1.** Typical Fourier spectrum of an ECG of a healthy heart. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

The cluster with the largest set of ECG is called the base cluster. For all other clusters, the measure of cluster deviation from the base as Euclidean distance between the corresponding averaged forms is introduced.



**Figure 2.** Typical Fourier spectrum of an ECG of a patient with advanced ischemic process. Horiz. Axis-frequency in Hz, vertical. Axis - potential in mV.

The developed ischemic process in accordance with the hypothesis is the formation of an autonomous dynamic system in the myocardium, which has oscillatory parameters different from the parameters of oscillations in the normal tissue of the myocardium. In particular, the main frequencies of the ECG spectrum of a normally functioning heart are in the range of 1-12 Hz, while the spectrum of the ischemic process is in the range of 14-20 Hz and can be dynamically changed.

On the basis of the characteristic features of the individual clusters obtained and the data on patient's pathologies, clusters similar in characteristics to the shape of the spectrum are identified and displayed in separate groups on pathologies. Thus, the clustering method can be used to diagnose ischemia, determine the cardiological status of healthy people.

## 7. Conclusion

The article discusses the possibility of storage and processing of cardiac information for mathematical analysis using cloud technologies.

During the study, the possibilities of analysis and monitoring of medical data, as well as forecasting in the tasks of mathematical analysis of medical data using cloud technologies were reviewed.

The mathematical modeling of the heart by using the decomposition of Fermi-Pasta-Ulam is carried out. To describe the mathematical model of the heart, we present a system of connected cells of myocytes, representing separate self-oscillating degrees of freedom, described by a system of coupled nonlinear differential equations of Van der Pol.

Cluster analysis is carried out on the basis of search of similar forms of Fourier spectrum obtained by modeling the heart FPU.

The main result of this work is the implementation of processing and storage of cardiograms (any number of leads), photoplethysmograms (pulse wave), echocardiograms, seismocardiograms for further calculations in the cloud and the organization of access to the results.

Thus, the use of the cluster approach can be useful for the diagnosis of various degrees of ischemia, determining the cardiological status of healthy people with an assessment of its age-related changes. In addition, the proposed cluster method will assess the effect of the therapy in dynamics, and change it in case of inefficiency.

## 8. References

- [1] Andreu J, Poon C, Merrifield R D and Yang G-Z 2015 Big data for health care *IEEE Jour. of Biomedical and Health Inform.* **19** 1193-208
- [2] Braunwald E 1997 *Heart disease: textbook of cardiovascular medicine* (Philadelphia: W. B. Saunders Co.) p 108
- [3] Fermi E, Pasta J and Ulam S 1955 *Studies of nonlinear problems* (Collected Papers of E. Fermi University of Chicago Press) p 978
- [4] Zabusky N and Kruskal M 1965 Interaction of solitons in a collisionless plasma and the recurrence of initial states *Phys. Rev. Lett.* **15** 240-243
- [5] Van der Pol B 1926 On relaxation-oscillations *The London, Edinburgh and Dublin Phil. Mag. & J. of Sci.* **2** 978-992
- [6] Novopashin M, Shmid A and Berezin A 2017 Fermi-Pasta-Ulam auto recurrence in the description of the electrical activity of the heart *J. Medical Hypothesis* **101** 12-16
- [7] Berezin A 2004 Resonant interaction between relapses Fermi-pasta-Ulam *Bulletin of the Physics Institute FIAN* **3** 13
- [8] Ezhkov A, Isakov R, Sountsova O and Sadovskiy S 2016 Connection between heart control and lungs pathology based on the analysis of an electrographic signal recorded by the cardiac monitor cardioqvark *Research Journal of Pharmaceutical, Biological and Chemical Sciences* **7** 3137-3142
- [9] Rubanik V 1969 *Oscillations of quasilinear systems with a time lag* (Moscow: Science) p 288
- [10] Moyer V 2002 Screening for coronary heart disease with electrocardiography: recommendation of the us preventive services task force *Annals of Internal Medicine* 157
- [11] Novopashin M, Schmid A and Berezin A 2017 Forrester's concept in modeling heart dynamics in *IOSR Journal of Computer Engineering IOSR* **3** 113-121
- [12] Ikezi H 1978 Solutions in plazma *Solitons in Action* 163-168
- [13] Lichtenberg A and Lieberman M 1992 Regular and chaotic dynamics *Applied Mathematical Sciences* **38**
- [14] Longren K 1978 Solitons in nonlinear electric lines *Solitons in Action* 138-162