Big data analytics with graphical techniques applied on sensors data

M L. Oroian-Boca
“1 Decembrie 1918” University of Alba Iulia, Faculty of Exact Science and Engineering, Department of Exact Science and Engineering, Alba Iulia, Gabriel Bethlen street, no.5, Romania

E-mail: loredana_boca1@yahoo.com

Abstract. The term “Big Data” is now a popular way to refer to massive digital information available in both structured and unstructured form integrated from multiple, diverse, dynamic sources of information. Big data can be used to solve a variety of problems with significant cost reduction cost. Analysis and processing of large data sets represent a significant challenge.

Massive data sets are collected and studied in numerous domains, from engineering sciences to social networks, biomolecular research, commerce, and security. Extracting valuable information from big data requires innovative approaches that efficiently process large amounts of data as well as handle and, moreover, utilize their big data analytics structure can be used in electronics industry too. In this study, I created a graphic tool in LabVIEW for the analyses of ECG signals for 3 types of subjects, in the same environment take in case the age of each subject and the influence of the environment on the physiological state of each one.

The environmental conditions varied and the variations in the human body have been assessed in three situations: studying with music, studying without music and in relaxing state. LabVIEW proved to be an efficient tool to analyse and compare the produced signals in various time spans and also to compare the ECG signals for the investigated subjects.

1. Introduction

Most efforts towards analysing Big Data assume data parallel applications and handle the large volumes of data. However, Big Data is characterized in general by the 4V’s – Volume, Variety, Velocity and Veracity. Efficiently handling large amounts of data is a big challenge to current computational resources.

There exist two common sources of big data, collective gathering and individual generation. Examples of collective gathering big data are smart city data, conditions monitoring data, and earth observation data [1]. In general, collective gathering data are obtained with sampling strategies, and its data quality is high.

Examples of individual generation data are electronic business data, social media data on the Internet, and crowd sourcing data [2]. Individual generation data is obtained in more freedom, and its usability is relatively low. The big data technology arises from the rapid development of the computer and communication technology, which evenly causes the paradigm transition of human cognition about our living world. On the one hand, the computer technology, particularly the digital and intelligent technology, is much advanced at the stages of data sampling, storage management, data computation, and data communication.
The essential of big data analytics is the structural analysis of big data in an optimal criterion of physics, computation, and human cognition [3]. Thus, it is safe to assume that technology advancements and practices such as big-data and advanced analytics will likely have a definitive impact in addressing some of these challenges it is shown in figure 1.

Electrocardiogram - ECG has proved out to be very useful human heart beat and in some case diagnosis human behaviour. The design ECG acquisition system is difficult environment is not always available for analysis of these systems. The existing systems are more expensive. The use of instrumentation system based on National LabVIEW software makes the system cost be utilized as a test bench for the study the laboratory level.

Electrocardiography is a method of monitoring and recording the electric currents generated during the alternating contractions of the atria and ventricles of the heart. The device used to monitor and record these signals is an electrocardiogram, more commonly referred to as an ECG.

The ECG records the electrical activity of the heart where each heart beat is displayed as a series of electrical waves characterized by peaks and valleys.

Any ECG gives two kinds of information.
1) The duration of the electrical wave crossing the heart which in turn decides whether the electrical activity is normal or slow or irregular.
2) The amount of electrical activity passing through the heart muscle which enables to find whether the parts of the heart are too large or overworked [4].
Figure 2. Example of ECG waves and intervals, [5].

Normally the frequency range of ECG signal is 0.05-100 Hz and its dynamic range 1-10 mV. The ECG signal is characterized by five the letters P, Q, R, S, T. In some cases, also it uses another peak called U. The performance of ECG analysing system depends mainly on the accurate QRS complex, as well as T or P-waves.

The P-wave represents the activation of the upper chambers of the heart, the atria, while the QRS complex excitation and T-wave represent the excitation of the ventricles or the lower chamber of the heart. The detection of QRS complex is the most important task in automatic ECG signal analysis.

Once the QRS complex has been identified a more detailed examination of ECG signal including the heart rate, the ST segment can be performed. Normal ECG with standard values is shown in figure 2. Most software packages are either general purpose programming languages, which do not usually contain any signal processing libraries, or dedicated turn-key software that perform only a single task (i.e. acquisition).

Few addresses all the requirements of a measurement system, including analysis. Unlike software development tools designed only for data acquisition or only signal processing, LabVIEW was developed from the beginning to provide a completely-integrated solution, so that users can simultaneously acquire and analyse data in a single environment [5, 6].

2. Big data description

Big data can be used to find solutions for different complex problems through analytics such as:

- Predictive Maintenance (PDM): An application can acquire data from a Unit Under Test (UUT) at regular intervals and constantly analyse to detect an acute failure or degradation of performance over time.
- Performance Analysis: Metrics can be used along with data acquired in real-time to create values representative of performance which can then be used to identify cost. This data could also be used to create a baseline to quantify the effect of process changes.
- Simulation Models: A simple model can be created by playing back acquired data or a more sophisticated model can be created using acquired data to create a statistically similar model. This model can be used to create a control algorithm without hardware and/or avoiding additional edge testing that could be destructive and/or expensive [7].
A typical system of Virtual Instrumentation (VI) is constituted by a set of devices capable of communicating between them and for a program that controls the mentioned communication. The device that controls the system of virtual instrumentation is a computer, whereas other devices are measurement devices and other equipments.

From the information loaded by the different devices we can change the conditions of the test, modifying parameters of the measurement instruments. In an environment as the described one, the current trend is that specialized software controls the system, coordinating the functioning of the different elements.

One of these programs software is LabVIEW of National Instruments. LabVIEW allows to load, to analyse and to monitor the information inside a graphical programming environment in which there are assembled objects called virtual instruments to form the program of application with which the user will interact.

In addition, LabVIEW allows the information representation in interactive panels that work as if it was a real instrumentation, and it allows multiple options of information treatment, as his storage on disc and to share them in network or with other applications [5, 6, 8].

Each sensor measures various experimental parameters and the recorded values are stored in Coma Separated Values (CSV) or Excel files. Such kind of files is difficult to be managed because each of them contains millions of data (usually around 2 million of values). Also, the research was made in different case study and for different persons. According to this situation, in a period of time a sensor has stored ~50000 records [8]. Figure 3 represent the average number of values in every file ~2 million; resulting in a very large Excel file (~30-150 MB)

![Figure 3. Types of Excel Data collected from sensors.](image-url)
3. Case study

In this paper I propose a method to view and analyse big data in a case study in which the methodology explores the hypothesis that a person’s physiological state can be assessed and improved by adequate sensorial stimulation and further on to relate it to its emotional state. In this way it would be possible to infer a person’s emotional state by considering physiological measurements, as we know that emotional states are not possible to classify by direct assessment. This case study was done on different subjects with different age in the same conditions: study with music, study without music and in relaxing time.

LabVIEW with its signal processing capabilities provides a robust and efficient environment for resolving ECG signal processing problems. LabVIEW has powerful tools for denoising, analysing, and extracting ECG signals easily and conveniently. [6].

These tools can be also used in other biomedical signal processing applications such as Magnetic Resonance Imaging (MRI) and Electroencephalography (EEG).

Now in LabVIEW Biomedical Toolkit, several VIs is provided for ECG signal analysis. Besides, it also contains an ECG Feature Extraction application to extract ECG features more conveniently, [5]. ECG is a non-invasive, convenient, and reliable method of measuring the electric characteristics of the heart. I have conducted physiological measurements on a group of 3 students with different age namely electrocardiography (ECG) to analyse the influence of music on the attention state of the student. The students which are taking in case have different age, one of them around 20 years old and other around 30 years old.

LabVIEW with its signal processing capabilities provides a robust and efficient environment for resolving ECG signal processing problems, [8]. Spectral analysis of HRV was contemplated in 3 young male subjects at different ages during their engagement in PC based Cognitive Tasks consisting of Auditory Tone Discrimination, Working Memory and Continuous Performance (Vigilance) Tasks, each with two or more levels of memory load/time compression. Manipulation of cognitive demands in these task variants was objectively defined.

Figure 4 shows the view of date from excel files used in LabVIEW. First array Data 1 contain the ECG data study with music, the second array Data 2 contain the ECG data study without music and last Data 3 contain the ECG data in relaxing time.

**Figure 4.** Big Data from Excel files uploaded in LabVIEW.
Figure 5 shows the view of date from excel files used in LabVIEW, apply on 3 types of subject one on 33 years old, one 22 years old and one on 24 years old, in the same condition study with music.

Figure 6. Graphic analyses for each case study.

7. Graphic analyses for each case study in the same condition study with music apply on different subjects.

Figure 6 shows the graphic analysis for each case study. I took into consideration one student and I stored ECG date in 3 situations:
- Situation 1 in which the student is studying without music;
- Situation 2, when the student is studying with background music;
- Situation 3, when the student is in a relaxing state. We have selected 500 representative data for graphical analytics.
Figure 7 shows 3 types of subject one on 33 years old, one 22 years old and one on 24 years old, in the same condition study with music.

The block diagram of this application is presented in figure 8. The program block diagrams show the objects include terminals, subVIs, functions, constants, structures, and wires, which transfer data among other block diagram objects and after will be show in front panel.

4. Results
Results of the study shows in figure 9 disclosed the susceptibility of certain spectral components of HRV to cognition especially when locus of the load was on working memory. The effect was demonstrably independent of possible entrainment of power spectrum of HRV with the stimulus frequency and/or changes in breathing pattern.

After the analyses I obtain some difference between subjects take in case same condition study with music at different ages.

Modulation of HRV during cognition seemed to be correlational. The ECG signal’s amplitude is quite large.

Figure 8. Application block diagram.

Figure 9. Graphic comparison for tree subject on different age in the same condition of study.
5. Conclusions
In this study, I created a graphic tool in LabVIEW for the analyses of ECG signals of three subjects and the influence of the age on the physiological state of each subject in the same condition. The age varied and the variations in one situation: studying with music.

LabVIEW proved to be an efficient tool to analyse and compare the produced signals in various time spans and compare the ECG signals for the investigated subjects.

After the analyses I obtain some difference between subjects take in case same condition study with music at different ages, means that the attention and study focus is more develop at young subjects then older one.

6. References
[1] Li D R, Yao Y and Shao Z F 2014 Geomatics and Information Science of Wuhan University 39(6) 630–640
[2] Shaw S L and Fang Z X 2014 Geomatics and Information Science of Wuhan University 39(6) 667–670
[3] Hong Shu 2016 Geo-spatial Information Science 19(2) 119-128
[4] Sandryhaila A, Jose M and Moura F 2014 IEEE Signal Processing Magazine 31(5)
[5] Deshmukh A and Yogendra Gandole 2014 International Journal of Science and Research (IJSR) 3 2319-7064
[6] LabVIEW for ECG Signal Processing, http://www.ni.com/tutorial/6349/en/, accessed February 20th, 2019
[7] https://tech-talk.org/2015/01/23/big-data-infographics-images/ accessed at February 21st 2019
[8] http://www.ni.com/tutorial/6349/en/#toc3 accessed at February 21st 2019

Acknowledgments
The research leading to these results has received funding from the European Union ERASMUS+ Program under grant agreement Higher Education – International Capacity Building – ACACIA – (561754-EPP-1-2015-1-COEPKA2-CBHE-JP).