|  |  |
| --- | --- |
|  | FORECASTING THE ONSET OF ACUTE MYOCARDIAL INFARCTION USING HIDDEN PREDICTIVE VALUES OF ECG SIGNAL  Research 3A Final Report |
|  | Advisor: PhD Le Quoc Trung |
|  |  |
|  | Student: Pham Khoi Nguyen🞂BEBEIU13051 🞂 12/29/2015 |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

# ABSTRACT:

Electrocardiogram (ECG) has currently been considered as a Golden test for approving a patient to the Cardiology Department after the patient developed an acute chest pain because the technique is cheap, fast and reliable. ECG measures the electrical activity of the heart and therefore provides significant information regarding the severity of cardio vascular diseases. One of the most important application of ECG is detecting the presence of Acute Myocardial Infarction (MI) and localizing the MI side. While the current application of ECG benefits greatly in the diagnosis field, the emerging technology and methodology regarding database analysis and statistic, for example: Data Mining, Neural Network… yield an even more promising application of using ECG signal not only in the field of Medical Diagnosis, but also of Predictive Medicine. This research aims to identify the hidden yet interpretable dynamic values of ECG signal in order to gain a deeper understand about dynamical processes of the electrical activity of the heart that eventually lead to Acute Myocardial Infarction (AMI). The knowledge can then be applied to create a model that could forecast the onset of AMI before clinical symptoms occur.

# INTRODUCTION

Acute Myocardial Infarction (AMI), also known as Heart Attack, is a disease caused by insufficiency of blood supply to the heart’s tissue. Generally, heart’s tissue is supported by a system of blood vessels. When these blood vessels suffer from Coronary Artery Disease (CAD) – the constriction of the artery that obstruct blood flow by the formation of fat and cholesterol beneath the vessel’s inner wall, some part of the heart does not receive enough blood supply. This phenomenon, if left untreated for a period of time, can eventually lead to cells death. The condition when some region of the heart died because of the derivation of blood supply and cannot function normally is called Myocardial Infarction.

Acute Myocardial Infarction is the leading cause of death during patients’ hospitalization in the United State. It has been reported that each year in the America, 1.1 millions of people suffer from Myocardial Infarction and haft of them get an acute attack [4]. In Vietnam, the number of patient who suffered from AMI tends to increase drastically during the last 20 years (from 1980 to 2000): from 1980 to 1990 there was 108 patients, from 1991 to 1995 there was 92 patients and from 2000 to 2001 the number escalated quickly: more than 1.500 patients [4] and 17.4% (261 patients) died [4]. Most of these patients are elderly (>65 years old) whose biological characteristics of coronary artery make them more susceptible to AMI.

Treatments for Acute Myocardial Infarction: Stenting and Coronary Bypass Graft for example, are extremely time – consuming and expensive. Therefore, the current diagnosis techniques strives to detect AMI before tissue death occurs. Some powerful medical techniques include: Cardiac MRI, CT Angiography and Echocardiography. These techniques yield considerable advantages: high accuracy, clear image resolution and many other reliable medical diagnostic values. However, when it comes to the need of real – time monitoring, these techniques are not considered as appropriate approaches dues to their high cost, time – consuming preparation procedures and cumbersome devices. Not only that, the technician who are qualified enough to perform these diagnosis are not always available in many hospital, especially medical centers from rural areas. Therefore, this research aims to develop a new technique for real – time monitoring and forecasting the occurrence of AMI using a famous but also very reliable signal: the ECG.

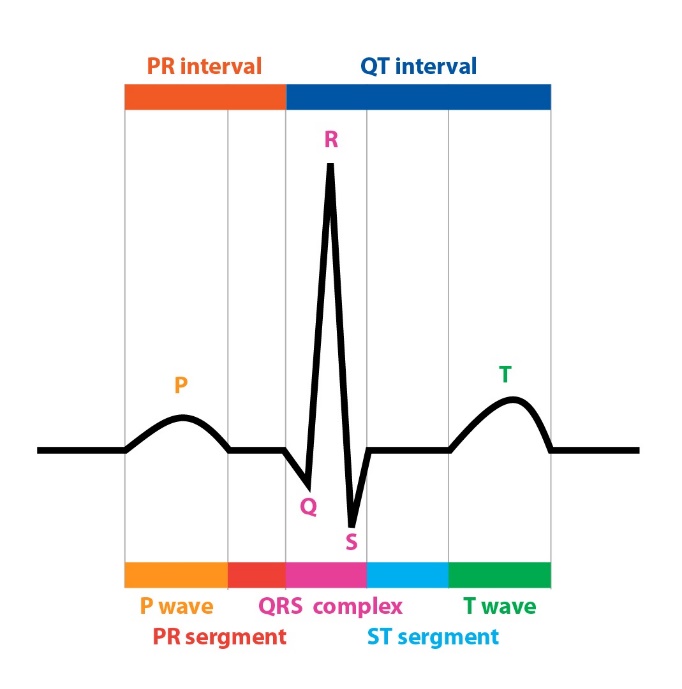
ECG, or Electrocardiography, measures the electrical activity of the heart during each heart beats. Current medical diagnosis using ECG makes use of the wave form and magnitude of the signal and delivers considerable amount of information about the heart from it. By looking at some specific segments of the signal: P wave, T wave, the presence of Q wave and ST segment, detection of Acute Myocardial Infarction can be achieved.

For the above reasons, ECG measurement has been used as a standard procedure for approving patients to the Heart Disease Department in almost hospitals around the world because it is cheap, fast and reliable. Then, further diagnosis techniques using more advanced devices are applied if necessary. However, the prognostic value of ECG for Acute Myocardial Infarction and the role of using ECG for forecasting the occurrence of AMI is still a an uncultured field but yet extremely profitable if it was discovered. The model can help elder people prevent the occurrence of AMI or make immediate responses to sudden cardiac attack.

# LITERATURE REVIEW

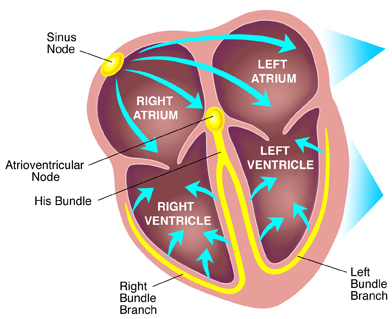
## The electrical activity of the heart

Figure 1.1 represents a typical example of the electrical activity of the heart. ECG (Electrocardiogram) is a sequence of wave forms that manifests the dynamical activity of the heart during consecutive heart beats. Each part of the wave form corresponds to a specific function in a specific region of the heart. Our heart consists of 4 champers: 2 atrial receiving blood from the body and 2 ventricle that receive blood from the atrial and pump back to the body. The pumping activity of these champers are described as the peak value in the ECG wave form, where the first peak (P wave) represents for the depolarization and contraction of the atrial, and the second peak (R wave) represents for the depolarization and contraction of the ventricle. The final peak (T wave) manifests the repolarization, or relaxation, of the ventricle thus preparing the heart for the next heart beats.

When analyzing ECG signal, the combination interpretation of waves together with the interval between these waves can yield reliable information about the overall activity of the heart. Each normal heart beat consists of a P wave, a QRS complex and a T wave and the corresponding interval in between. In figure 1.2, the electrical current that flows through the heart in each consecutive heart beat is described, providing the physiology perspective of the intervals. 

**Figure 1.1 a normal ECG wave form**

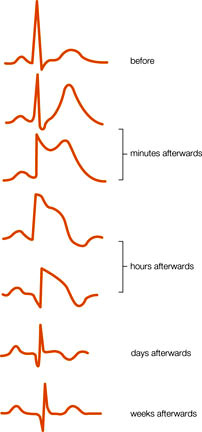
First an action potential is generated at the SA node, producing an electrical current that flows to the atrial. This electrical current stimulates the contraction of the atrial, resulting in the formation of P wave in the ECG signal. The current then flow through the atrial to the AV node, where it then travels to the left and right branch of the ventricle by conductive fibers known as left bundle branch and right bundle branch. The arrival current then stimulates the contraction on both the left and right ventricle, representing the R peaks in QRS complex. It is also important to mention that during this interval, the simultaneous contraction of the ventricle and relaxation of the atrial contribute to the complexity of the QRS pattern. After this point, ST segment represents for the time delay between ventricle contraction and relaxation, where all the ventricular tissues, after contraction, tend to bounce back to normal in order to be ready for the repolarization. This is the most important interval for spotting any damage to the ventricle tissues because the electrical current developed by any injury or inflammation to the ventricle appears clearly during this isoelectric process. Finally, the T wave occurs as the result of ventricle repolarization, or relaxation.

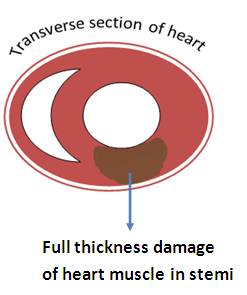
****

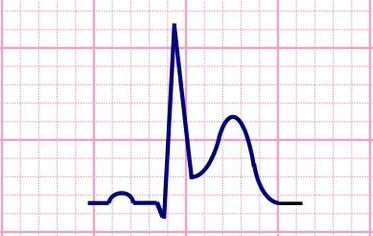
**Figure 1.2: Electrical conduction system of the heart**

## Manifestation of AMI in ECG signal

As previously mentioned, the ST segment is the most valuable part of the ECG waveform to identify any damage to the ventricle, where most AMI occur. Beside ST segment, other valuables waves that also need attention is the T wave and Q wave. Acute Myocardial Infarction manifestation in ECG signal is a sequence of changes in the T wave, ST segment and Q wave as described in figure 2.1. Firstly, the constricted blood vessels prevents or decreases blood flow to some specific region of the heart. Atrial tend to have larger coronary vessels, therefore it is less susceptible to blood derivation. However, the ventricle is supported by a really complex vessel systems, some is big (Bundle of His) and some is really small (Purkinje Fibers), therefore the restriction of blood supply is more viable, thus making the ventricle more susceptible to blood derivation. When the tissue is lack of blood supply, the Ischemic Event occurs, resulting in the injury or inflammation in some region of the heart. These injury generates addition electrical current that can be detected during isoelectric process of the heart: the ST segment. During this phase, T wave first becomes peaked and then ST changes occur. If the ST segment elevates, the myocardium is interpreted as having full thickness damage of the heart muscle (Figure 2.2a). If the ST segment depresses, then the traverse damage can be the cause (Figure 2.2b).

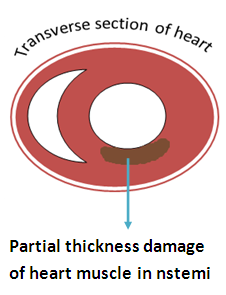
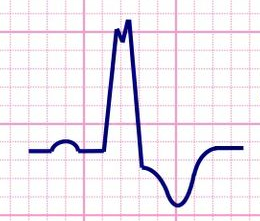


**Figure 2.1: Dynamical changes of ECG waveform during the formation of AMI**



**Figure 2.2a: ST segment Elevation**

Finally if the disease is untreated for a long period of time, tissues death occurs. The formation of a pathological Q wave (larger and more negative Q wave) develops. This is also the final stage of Acute Myocardial Infarction.



**Figure 2.2b: ST segment depression**

In conclusion, if the ST segment are elevated representing tissue injury, the phenomenon is categorized as ST Segment Elevation Acute Myocardial Infarction (STEMI). If the ST are horizontally normal or depressed representing ischemic event or tissue injury, it is categorized as Non – ST Segment Elevation Myocardial Infarction (NSTEMI).

## Data Mining and Hidden Predictive Value of ECG

Beside the medical point of view, the advancement in Information Technology has made the birth of Big Data, where thousands of databases in a tremendous variability of application fields are easily found and accessible via the internet. Despite the tremendous amount of data available, little useful information has been derived from these database that can be turned into practical knowledge. Within the condition where the blossom of data has become more available, Data Mining plays an essential role in how we make use of these advantages. Data Mining is the process in which useful information that can be turned into practical knowledge is being observed and taken out of the Big Data set.

Regarding Medical Application, medical data can be analyzed to produces significant information that can be used for Diagnosis and Predictive purposes. In our research interest, we strive to integrate Data Mining into the research methodology in order to study and understand the hidden dynamical process that lies within the ECG signal itself. To be more elaborate, large amount of ECG signals from various databases will be taken into account, where different analysis techniques are applied to find the underlying features that eventually lead to Acute Myocardial Infarction. These features are referred as the Hidden Predictive Value of ECG.

Section I.1 and I.2 has described ECG signal in a physiological perspective. Now we are looking at this signal from another different perspective: Statistics.

While the signal is observed quite as periodic processes since the consecutive waveforms tend to repeat over the time, the overall signal is still considered as a chaotic, nondeterministic signal that is also susceptible to noise. Therefore in order to adequately analyze ECG signal using probability and statistic technique, several requirements have to be taken into account:

1. The technique must be able to express the randomness of the signal.
2. While the data is chaotic and nondeterministic, periodicity must be observable.
3. Transition Probability of one stage to another has to correspond with physiological meanings.

During the analysis process to withdrawn the underlying ECG features, it is important to turn ECG signal from Time Series Domain into different other domains such as: Frequency Domain and Phase Space Domain.

#### Time Series domain

In Time Series Domain, ECG signal is a sequence of measurement that takes place over the time. This domain gives the most basic information about the electrical activity of the heart through different cardiac circle and information regarding these activities is manifested in the shape of the waveform. In order words, this domain allow us to withdraw information about the physiological activity of the heart over the time by interpreting the shape of the waveform. However, this technique depends heavily in the visual ability to spot out the unusual patterns of the signal, not taken into account the dynamical changes that caused the formation of these abnormalities. Although applying this visual technique in detecting the abnormalities within these signal yields appropriate information for detection of AMI, it is still not an appropriate approach for prediction because at this point, clinical symptoms have already occurred.

#### Frequency Domain

The process of transforming ECG signal from time series domain into frequency domain is known as the Fourier Transform where the equation is described as following:

\hat{f}(\xi) = \int_{-\infty}^\infty f(x)\ e^{- 2\pi i x \xi}\,dx,

In this equation, *f(x)* is the modelling function in Time Series Domain. The transformation allows the breakdown of the original signal into its frequency components versus the signal amplitude.

In reality where the modelling function of the signal is not available or too complicated to calculate, Fast Fourier Transform (FFT) is applied directly to a Time Series data set and turns it into frequency components. In this domain, information derived is the frequency of a specific signal within the original signal and usually this technique is applied to withdraw the desired signal from the noisy output. After the range of frequency of interest has been selected and further processed, the inversion technique is applied to return the signal into the original Time Series domain:

f(x) = \int_{-\infty}^\infty \hat f(\xi)\ e^{2 \pi i \xi x}\,d\xi,

After the reversed transformation, visual technique is the applied to spot out the unusual waveform that corresponds with AMI.

#### Phase Space Domain

Phase space domain represents for every possible states in which the signal can be. In this domain, each factors that contribute to the formation of the signal is categorized as a parameter, and each parameter is described as an imaginary axis in the phase space diagram. The state of the system is presented as a unique point within that diagram.

The transformation of a signal into state space domain is particularly useful for analyzing the periodicity of the signal. Therefore, in our research methodology, we perform state space analysis in order to understand what truly contributes to the dynamical process of the heart that making the signal repeats over the time. However, in order to perform phase space transformation, the state equation of the system must be known. This is not usually the case because as previously discussed, the model that represent ECG signal is nondeterministic. Therefore the state equation is not available or too complicated to compute. In this case we perform a probabilistic technique that works in a similar manner to phase space method. The method is discussed below.

## Hidden Markov Model

Hidden Markov Model (HMM) is a statistical method in which the system being analyzed is assumed to have different states that produce a sequence of different outputs. In our particular case, the output observed (ECG signal) depends on different states of the system, in which each state corresponds to a particular disease condition. The transition from one state (disease state) to another is known as the transition probability. This can be inferred from analyzing relevant cases in which the dynamical values has already been recorded in the database.

Now we are going to look deeper at this model with our purposes and perspective. Each HMM model consists of 4 parameters:

(Q, π, N, M)

In which Q = {qi | i: 1 🡪 n}: Total states of the system. In the simplest model to predict AMI, we propose a model with a total of 6 different states presenting for 6 different states of AMI:

{q1 = normal ECG, q2 = peaked T wave, q3 = ST elevation, q4 = ST depression, q5 = Inverted T wave, q6 = Q pathological)

Π = {πi | i: 1 🡪 n}: initial probability to have qi at time t = 0. The meaning of this parameter is that how randomly a state will suddenly develop.

N = {nij | i = 1 🡪 n, j = 1 🡪 n}: transition probability from one state to another. Because we have 6 different states, N is a 6x6 matrix that represents for every different transitions possible. The values within this matrix can be sensed from analyzing the relevant data set. However, these values are expected to be a very complex equation which depends not only in time but also on the previous ECG signal itself. Therefore determining these transition values is the heart of the research. We propose using regression model to evaluate each of these transition probability values.

M = {mi(x), i = 1 🡪 n}: probabilistic output, meaning that if we have the probabilities calculated, how likely these values will contribute to the changes in the output ECG. The final model will be discussed in part III.

## Regression

Regression is the statistical technique applied to study the relationship between variables and output. In time series analysis, a simple regression model usually involves calculating the output basing on the variable time.

Y = A + BX + e

In which Y is the output, A is the intersection point, B is the slope describes how time X contributes to output Y, and e is Gaussian white noise.

However, in other to understand the dynamical process of the signal, it is viable to describe an output basing on the previous measurements together with time. In other words, in our research methodology we integrate a multivariable auto regression model which is described as:

Describing this in words, the output of the model is a weighted sum of some of the previous output. We propose using this technique to calculate the transition probability and probabilistic output of the HMM.

# METHOD

Combine the knowledge about the physiology of the disease together with statistical technique for analyzing, we propose the prediction model as described:

Real time ECG input

# IMPLEMENTATION PROCEDURE

Step 1: Specifying the appropriate database for training the model

Firstly, an extensive effort will be made to specify which ECG database on PHYSIONET will be cultured. The data selected to train the model must contain real time measurement of ECG signals for a particular long period of time (from 24 to 48 hours) even before an acute chest pain happens. Together with this criteria, the final medical condition of the subjects must also be documented: the occurrence of AMI and mortality rate.

Step 2: Finding the common ECG patterns and applying weighting factors

In this step, the transaction phase when normal ECG slowly develop into the acute phase is investigated. Common patterns must be seen across several patients when the normal ECG develops into abnormal one that signifies the early presence of acute myocardial infarction. In addition, a weighting factor is applied to these patterns, representing the likely hood that whether or not this pattern is a manifestation of AMI [8]. After performing calculation and projection, the underlying features is announced.

Step 3: Developing the algorithm and the prediction model

At this stage, non – linear mathematical and statistical analysis is applied to create the forecasting model of AMI [9]. This model combines information derived from the previous stage and presents them as a system of non – linear differential equation for long term or short term prediction of AMI.

Step 4: Validating the accuracy of the algorithm using other sources of data

This step involves the process of comparing the result obtained by the model’s algorithm with the actual documented cases. The sensitivity of the model is then recorded and further improvement is applied if necessary.

# RESULTS:

# REFERENCES

[1]. Neil J Weissman, Bryan Ristow, Nelson B Schiller. “Role of echocardiography in acute myocardial infarction”.

[2]. Y Birnbaum, B J Drew. “The electrocardiogram in ST elevation acute myocardial infarction: correlation with coronary anatomy and prognosis”.

[3]. Jordan M Prutkin, MD, MHS, FHRS. “ECG tutorial: Myocardial ischemia and infarction”.

[4]. Tirza Springeling, Alexia Rossi, Adriaan Moelker, Robert-Jan M van Geuns. “Cardiovascular MRI in acute myocardial infarction”.

[5]. Paolo Raggi, MD; Tracy Q. Callister, MD; Bruce Cooil, PhD; Zuo-Xiang He, MD; Nicholas J. Lippolis, MD; Donald J. Russo, MD; Alan Zelinger, MD; John J. Mahmarian, MD. “Identification of Patients at Increased Risk of First Unheralded Acute Myocardial Infarction by Electron-Beam Computed Tomography”.

[6]. Journal of the American College of Cardiology. A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. “ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction”.

[7]. Chuyên ngành: Nội – Tim học. Luận án tiến sĩ y học. Nguyễn Văn Tân “Nghiên cứu sự khác biệt về lâm sang và cận lâm sang về điều trị nhồi máu cơ tim cấp ở bệnh nhân trên và dưới 65 tuổi”.

[8]. B. Taha, S. Reddy, T Aufderheide . “An Adaptive Fuzzy Model for ECG Interpretation”.

[9]. Li Sun, Yanping Lu, Member, IEEE, Kaitao Yang, and Shaozi Li. “ECG Analysis Using Multiple Instance Learning for Myocardial Infarction Detection”.

[10]. H. Michael Bolooki, Arman Askari. “Acute Myocardial Infarction”.