Capturing, Analysing and Managing ECG Sensor Data in Handheld Devices[1]

J. Rodríguez, A. Goñi and A. Illarramendi

University of the Basque Country (UPV/EHU)
LSI Department. Donostia-San Sebastián. Spain
http://siul02.si.ehu.es

Abstract. The new advances in sensor technology, PDAs and wireless communication favour the development of monitoring systems. In this context we have developed MOLEC, a system that monitors people that suffer from heart arrhythmias. It carries out three tasks: a) it *captures* the digital data sent by different sensors and transforms them into a concrete format; b) it *analyses* the ECG data sent by the PDA sensors to locally detect anomalous situations as soon as possible; c) it *manages* ECG data in the following way: 1) normal ECG data are stored in files and are sent compressed to a hospital, within a certain time granularity; and 2) when anomalous ECG data are identified an alarm is sent to the hospital; and also these anomalous ECG are stored at the PDA database. Data in that database can be queried, locally or remotely, in order to know different aspects that can be related to anomalous situations.

1 Introduction

Innovation in the fields of PDA, wireless communication and vital parameter sensors enables the development of revolutionary monitoring systems, which strikingly improve the lifestyle of patients, offering them security even outside the hospital. Focusing on electrocardiogram (ECG) sensors, it is important to see that the new ECG monitoring systems outperform traditional holters.

The use of a holter consists in placing electrodes (leads) on the patients' chests; these leads are attached to the holter. After the patient is sent home and goes back to normal life, a tape records a continuous ECG for 24 or 48 hours. One or two days later, the holter is removed and the tape is analysed. The physician will see each of the patients' heart beats and if abnormal beats or heart arrhythmias occur during that period, they are identified by the physician [1]. Although this solution presents the advantage that patients can continue living a normal life in their houses, it also presents a serious drawback: if the patient suffers from a serious rhythm irregularity, the holter only records it, i.e. it does not react to it.

[1] This work was mainly supported by the University of the Basque Country, Diputación Foral de Gipuzkoa and the CICYT: Comisión Interministerial de Ciencias y Tecnología, Spain [TIC2001-0660].

In order to overcome the previous restriction, some new proposals have appeared that not only record ECGs but also react in some concrete situations. We classify these proposals in two major groups: commercial systems and research projects.

In the first group we include those commercial systems that use a mobile telephone unit or a PDA (Personal Digital Assistant) to acquire the ECG signal. Three or four metal electrodes are situated on the back of the standard cellular phone, which record the heart event. Then, the data are transmitted to the cardiac monitor centre situated at a hospital (e.g. Vitaphone [2]). Another alternative is to use PDAs. Companies like Ventracor [3] from Australia or Cardio Control [4] from the Netherlands have developed systems capable of storing ECGs directly in the PDA. Additional features like GSM/GPRS transmission to an analyzing unit are also being developed.

In the second group there are several research projects like @Home [5], TeleMediCare [6], or PhMon [7], whose aims are to build platforms for real time remote monitoring. These systems include wireless bio-sensors that measure vital parameters such as heart rate, blood pressure, insulin level, etc. The health monitoring system, carried by the patients, controls these sensors. The patient data recorded are sent to the hospital, where they are analysed.

All the previous systems are continuously sending ECGs to a hospital through a wireless communication network, where they are analysed. In spite of the advantages these kinds of systems provide in relation to holters, they still present main problems related to the fact that the analysis is not performed in the place where the signal is acquired. Therefore, there is a *loss of efficiency* in the use of the wireless network because normal ECGs are also sent (and wireless communications imply a high cost); and, in the case of the wireless network is not available (e.g. in a tunnel, in an elevator, etc.) at some moment, there might be a loss of ECG signal with the corresponding risk of *not detecting some anomalies*. Our proposal is to analyse the signal locally at the PDA.

In this paper we present a system called MOLEC: a PDA-based monitoring system that records user ECG signals in order to find arrhythmias, and in case of detecting them, it sends them to a hospital so that cardiologists can determine what to do with the user. The MOLEC system is separated into two subsystems, one situated at the hospital and another one situated in the PDA carried by the user.

MOLEC has been designed using a modular approach and applying distributed-object technology. Each one of the three main modules of the system is related to one task carried out during the monitoring process: to capture, to analyse and to manage ECG sensor data. Hence, the aims of the *Data Pre-processing Module* are to capture the data sent by the sensors, and to process them in order to generate a sequence of beats and associated information needed by the Decision Support Module. The *Decision Support Module* classifies the beats and the rhythms found in the ECG with the aim of finding abnormal situations. As for the third and last module, the *Database and Communication Module*, its task is to efficiently manage the data coming from the two previous modules.

Moreover, MOLEC system implies a great advance in the process of on-line monitoring of heart diseases, because it provides: 1) *efficiency*: premature detection of abnormalities by the monitoring system and optimization of wireless communications; 2) *local analysis:* even if wireless communication were unavailable, the signal could be analysed locally at the PDA; 3) *openness*: it can be integrated in hospitals that manage clinical data through the HL7 standard [9] (a representation of clinical documents); 4) *accessibility*: data recorded in the system can always be queried whether locally or remotely; 5) *simplicity*: making technical issues transparent to the users from the point of view of software and hardware components; 6) *adaptability*: possibility of working with different kinds of ECG sensors

In the rest of the paper, first we briefly explain the framework of the system and in sections 3, 4 and 5 we present the modules mentioned previously in more detail. At the end of the paper we show our conclusions.

2 Framework of the MOLEC System

As we have mentioned in the introduction, the goal of the MOLEC system is to facilitate the monitoring of the heart diseases. The following elements help in pursuing this goal: 1) the *ECG Sensors* carried by a patient, in order to register heart signals; 2) the *Sensor-PDA-Holter*, a mobile device carried by the user that acquires, processes and transmits ECG signals; 3) the *MOLEC Hospital*, which receives possible abnormalities sent by the Sensor-PDA-Holter and results of queries made by physician(see figure 1).

From left to right, figure 1 show: 1) The *ECG Sensors* acquire the electric impulse of the heart. These sensors are the 'intelligent' chips that communicate with the PDA through the bluetooth protocol. 2) We call the user mobile device the *Sensor-PDA-Holter* because it acquires the data signal in a PDA and works like a traditional holter. Besides, it detects abnormalities and notifies the hospital very quickly. The modules of the Sensor-PDA-Holter are the *Data Pre-processing Module*, the *Decision Support System Module* and the *Database and Communication Module*. 3) The *MOLEC Hospital* is a system that provides the Sensor-PDA-Holter with the user data in order to customis e it. It also receives the user's possible abnormalities. Hospitals would have to incorporate the MOLEC Hospital system into their administration system.

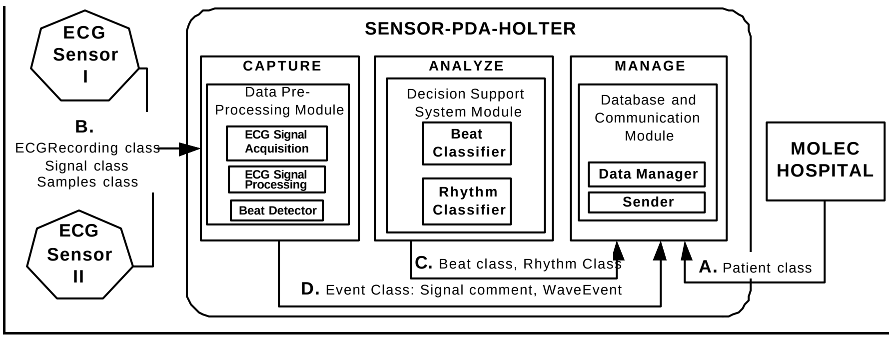

Fig. 1. MOLEC system

The data flows (see labels A, B, C and D in figure 1) represent the data exchanged among the previous modules. These data are objects of classes that express the data requirements needed to build a system that records ECG signals of users, processes them in order to identify the different beats, classifies each beat and also classifies the different beat rhythms with the aim of finding cardiac arrhythmias. These classes correspond to persistent classes that we implement by using database technology. They are also shown in the conceptual schema in figure 2.

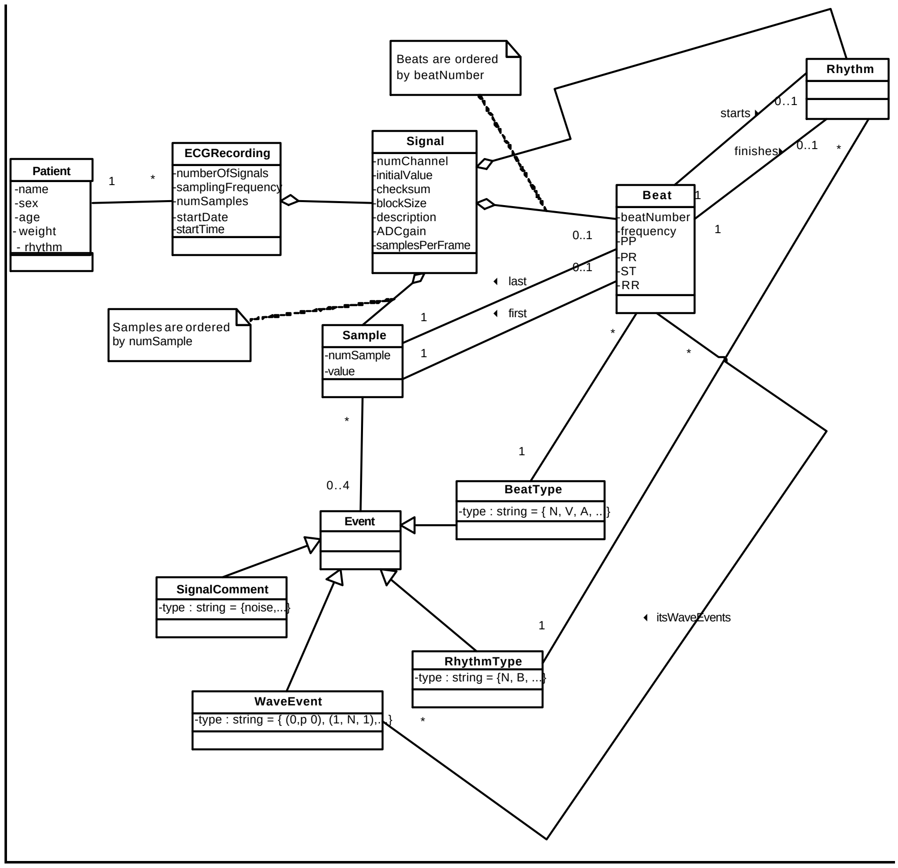

Fig. 2. Conceptual schema of MOLEC

In the following paragraphs, we explain the different classes present in figure 2 associated with the data flows that are used by the three modules that run at the PDA (i.e. Data Pre-processing Module, Decision Support Module and Database and Communication Module).

Data flow between MOLEC Hospital and Sensor-PDA-Holter (see label A in figure 1). The *Patient class* allows to store required data about users: name, sex, weight, age and typical heart parameters. These data must be stored in the hospital node, but, in every PDA, only the data corresponding to its user must be present. Those data are used to customize the PDA.

Data flow between ECG Sensor and Sensor-PDA-Holter (see label B in figure 1). The *ECGRecording class* stores different ECG recordings, each one belonging to a single user and with a sampling frequency, start time, start date, and number of samples. Each recording can consist of several signals (simultaneous recordings made in different channels). Each one of the signals is represented by a *Signal class* object. Each signal is seen as a sequence of samples, as a sequence of beats and as a sequence of rhythms. Each *Sample class* object has its sequence number and is associated with the signal object to which it belongs. These sample objects are obtained from the sensor and constitute the digital ECG signal.

Data flow inside Sensor-PDA-Holter (see labels C and D in figure 1). A sample object can be associated with four different events that are represented in the *Event class*. They are: 1) a *SignalComment class* object (for example, one that indicates if there is noise in the signal); 2) a *WaveEvent class* object (e.g. one that corresponds to a peak or to the limits of the P, QRS or T waves and also contains information about the type of T wave); 3) a *Rhythm class* object that indicates that a rhythm starts at that point of the signal; and 4) a *BeatType class* object that represents the type of beat at that point. All these event, beat and rhythm objects are obtained and attached to the sample events during the ECG signal processing by the beat detector, beat classifier and rhythm classifier. All the data corresponding to current ECG recordings are stored in the PDA; later, data can be stored in the hospital database if the physician wants to.

The rest of this article explains in detail the relationships between these classes and the PDA modules: Data Pre-Processing Module (section 3), Decision Support System Module (section 4) and Database and Communication Module (section 5).

3 Data Pre-processing Module

The data pre-processing module includes all steps necessary for a later correct analysis of the beats in the decision support system. These required steps include the acquisition of biological signals and the identification of each beat and its different parts.

The data pre-processing module (see Figure 3) consists of three submodules: 1) the *ECG Signal Acquisition*, which acquires the beat sequence from the ECG sensors and obtains the digital signals in a specific format; 2) the *ECG Signal Processing*, which receives digital signal segments and analyses them in order to determine the sequence of data that we call 'wave events' (the peaks and limits of P, QRS and T waves which are represented as points in figure 3); 3) the *Beat Detector* receives the wave events and identifies the RR, PR and QT intervals, the duration of the P, QRS complex and T waves, their frequency, and also determines the ST and PQ segments.

These three submodules convert the biological signal into a beat sequence with some important data such as wave events, intervals and frequencies, which will be used by the decision support module. We present each one of these submodules in the following subsections.

3.1 ECG Signal Acquisition

The ECG signal acquisition module, located at the PDA, manages the communication between the ECG sensors and the PDA. The electric impulses of the heart are obtained by the sensors and sent to the PDA. The information sent from the sensor to the PDA is a digital signal that has been transformed from biological signals. In order to completely understand this module, it is necessary to explain how an ECG sensor works, and how the data are translated to any format accepted by PhysioNet [11], in such a way that different ECG sensors may be adapted to MOLEC. In the following subsections we explain the submodules included in the ECG signal acquisition: the *ECG Sensor* and the *Sensor Data Adapter*.

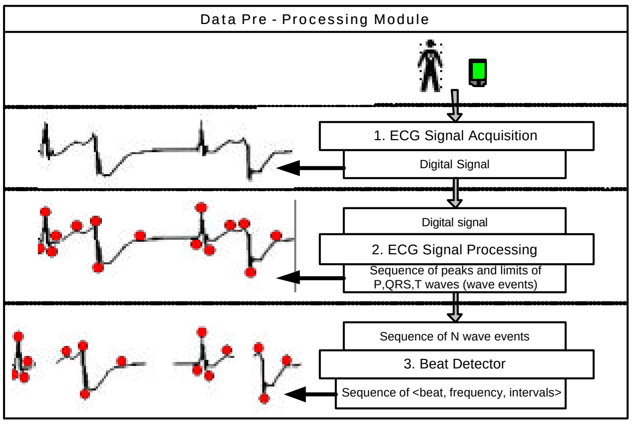

Fig. 3. Data pre-processing module

3.1.1 ECG Sensor. The electric impulse of the heart is acquired by the ECG sensors of the patient. These sensors are the intelligent chips that communicate with the PDA through the bluetooth protocol. The information to be sent is a digital signal that has been transformed from the analog signal.

Unfortunately, nowadays ECG sensor providers only sell their products with proprietary electrocardio analyser software. Therefore, in order to deal only with sensors, we have built an ECG sensor emulator that sends ECG data stored in a free downloadable database. The simulation process consists of dividing the data contained in the MIT-BIH database into bit sequences and sending those sequences periodically (see figure 4).

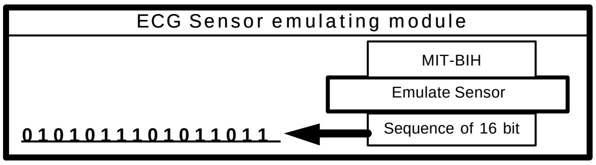

Fig. 4. ECG sensor emulating module

Through the development of this module, we have been able to provide the other modules with ECG signals, and to create a communication platform between the PDA and the sensors. Nowadays, there are some projects under development that try to build ECG wireless sensors, but they are not commercially available yet. We believe that with the architecture that we propose it would be possible to adapt our system in order to work with real ECG data sent by sensors.

3.1.2 Sensor Data Adapter and Visualisation. This module is located in the PDA and receives the signals sent by the ECG sensor. Next, the module translates the signal into a standard format that the whole system understands (Signal Class). The signal can be visualised on the MOLEC PDA screen (see figure 5) and is analysed in the next module, the ECG Signal Processing.

Fig. 5. ECG visualisation in Sensor-PDA-Holter of MOLEC

3.2 ECG Signal Processing

The ECG signal processing receives the digital signals and characterises the sequence of P, QRS and T waves associated with each beat. In order to read the ECG it is necessary to know the beginning, end, and the peaks of the waves that occur in a beat, because the absence, the format, the duration and other considerations allow some abnormalities to be identified.

0:00.819	279	(	beginning by P wave	0	0	0	0 corresponds
0:00.861	298	P	wave P	0	0	0	with the P
0:00.916	320	)	end P	0	0	0	wave
0:00.972	353	(	beginning by QRS	0	0	1	1 corresponds
0:01.025	370	N	wave QRS	0	0	0	with the
0:00.069	407	)	end QRS	0	0	1	QRS wave
0:01.219	424	(	beginning by T wave	0	0	2	corresponds with the T wave
0:01.300	455	t	wave T1	0	0	4	negative-positive
0:01.433	503	t	wave T2	0	0	4	negative-positive
0:01.488	523	)	end T	0	0	2	corresponds with the T wave

Fig. 6. The limits and peaks of a beat

In order to detect the limits and the peaks of the wave, we use ECGPUWAVE tool[2] [13], which is provided by PhysioNet, and implements the algorithm of on-line detection developed by Pan & Tompkins [14]. Its function is to identify the limits of the P, QRS and T waves in real time. This tool has been tested on different ECG

[2] Although ECGGPUWAVE was developed for a larger computer than a PDA, it runs in our prototype of MOLEC, that is implemented in a iPAQ 3970 with Bluetooth and with 48MB ROM, 64MB RAM, Linux Familiar 0.7, Java virtual machine 1.3.1 and MYSQL version 9.38.

databases: the CSE database [15], the MIT-BIH database [8] and the ST-T database [12]. ECGPUWAVE works with data in a specific format where the input data are a sequence of digital signals and the output data is displayed in the next figure.

In figure 6, the first column represents exactly when the signals were acquired. The second column is relative time in relation to acquisition frequency (e.g. 360 Hz). We refer to this in the rest of the paper as 'time unit' (t.u.). The last column represents the beginning and the end of P, QRS and T waves with open and closed parentheses respectively. These parentheses are associated with the following annotations (see the last column): 0 corresponds with the P wave, 1 with the QRS wave, and 2 with the T wave. The letters P and T denote the peaks of the P and T waves. The letter N denotes the peak of the QRS complex. The classification of the T wave is shown in the last column where 0 is normal, 1 inverted, 2 positive, 3 negative, 4 negative-positive and 5 positive-negative. To identify the limits and peaks of the beats, it is necessary to send this sequence of wave events to the Beat Detector.

3.3 Beat Detector

The beat detector receives the wave events and invokes specific automata that we have developed with the purpose of building up the beats. This purpose is achieved by scanning the wave events one by one and building up each beat. A beat usually begins with a P wave followed by a QRS complex and finishes with a T wave, but some of them may be missing and others may be repeated, especially when there are heart diseases. One typical format of the beat is shown in table 1.

Table 1. Two beats detected by the automata

ID	format	(	p	)	{	QRS	}	[	t1	t2	N_t	]
1	(p){N}[tt]	279	298	320	352	370	402	424	455	503	4	523
2	{N}[t]	-1	-1	-1	562	598	605	627	658	-1	0	712

In this table we see all events that could be present in one beat (row one), but the format of the beat could be any combination of these wave events (row two, in which the absence of the P wave is indicated by -1). In other words, the detection of the beat is not a trivial process (especially when abnormal beats occur in the user due to missing and/or repetition of waves that may correspond to different beats). Hence, the automata implement all combinations of those wave events and decide which is a beat and which is not a beat.

After, identifying the format of the beat, it is necessary to determine the ST and PQ segments and to calculate the RR, PR and QT intervals, as well as the duration of the P, QRS complex and T waves and the heart frequency. These values are necessary to identify the different anomalous behaviours that could be present in the user. In figure 7 we show the signal corresponding to row 1 in table 1. The portion of the ECG between the QRS complex and the T wave is called the "ST segment" and the portion between the P wave and the QRS complex is called the "PQ segment". The portion of the beat between the P wave and the beginning of the QRS complex is called the "PR interval" and the portion between the QRS complex and the end of the T wave is called the "T interval". The distance from a QRS complex to the next QRS complex is called the "RR interval".

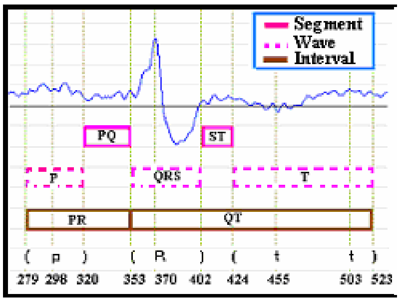

Fig. 7. Waves, segment, interval, limits and peak usually found in a beat

The duration of the waves is obtained by subtracting the end of a wave from its beginning. Looking at figure 7, it may be seen that the P wave of the beat began in time unit 279 and ended in time unit 320, i.e. it was 320-279 t.u. long. Do not forget that special values must be given when P waves, for example, have not been identified by the ECGUPWAVE tool (see table 1, row 2).

As mentioned above, this set of intervals, segments, waves, peaks, limits and format is necessary to classify one beat, so enabling the work of the next module, the Decision Support System.

4 Decision Support System Module

The decision support system is the module whose goal is twofold: to classify each beat and to classify each rhythm. The decision support module (see figure 8) receives a sequence of beats, their frequency and their intervals and classifies them as cardiac rhythms.

This module consists of two submodules: 1) the *Beat Classifier*, that classifies each beat thanks to a set of rules that we have defined with that purpose; 2) the *Rhythm Classifier*, which groups sequences of four beats and identifies the rhythm by using rules that we have defined and which reflect the properties of rhythms described in specialised literature [1]. The purpose of the decision support system is to classify each beat and rhythm in order to detect some abnormality. We present each one of these submodules in the following paragraphs.

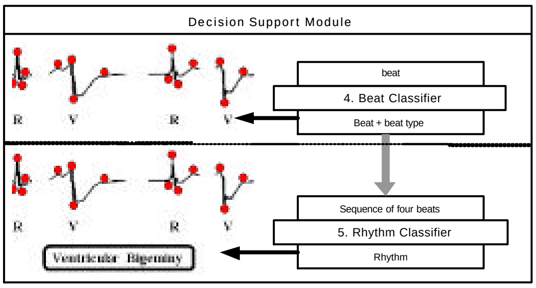

Fig. 8. Decision support system

4.1 Beat Classifier

The beat classifier receives the information related to the beats and applies a set of rules to classify them. In particular, we focus on methods that correctly classify the beat and rhythm types that PhysioNet provides.

PhysioNet contains a set of databases. One of them is the MIT-BIH database which contains 4,000 30-minute registers. However, we only have used the free 48 registers available (enumerated from 100-124 and from 200-234, with some missing in between). The first interval is a representation of typical clinical cases; in the second one, we find several complex anomalies like ventricular, supraventricular and nodal rhythms.

These registers are collected from men and woman between the ages of 23 and 89. The same instruments are used in all registers, which are first acquired analogically, and later transformed into digital signals, with a frequency of 360 Hz, using 11 bits and a resolution of approximately 5 mV. After the registers were collected, two independent cardiologists classified them, using tables 2 and 5, that show the different types of beats and rhythms, respectively.

Table 2. Table of beat type

F	Fusion of ventricular and normal beat	N	Normal beat
L	Left bundle branch block beat	E	Ventricular escape beat
R	Right bundle branch block beat	|	Isolated QRS-like artifact
j	Nodal (junctional) escape beat	"	Miss beat
f	Fusion of paced and normal beat	!	Ventricular flutter wave
A	Atrial premature beat	J	Nodal premature beat
a	Aberrated atrial premature beat	e	Atrial escape beat
V	Premature ventricular contraction	/	Paced beat
S	Supraventricular premature beat		

The process that we follow to find a set of rules that classify the beats is beyond of the scope of this article (see [10] for more details), but we can mention that a prototypical rule (table 3) is the following: a beat is normal if its R wave is between 0 and 26 time units, its frequency is between 0 and 63.16 t.u. and, finally, its RR interval is between 191 and 374 t.u. Once the beats are classified, it is necessary to classify the rhythms. The rhythms are classified using a set of beats and its components (segments, waves, frequency and intervals).

Table 3. Rule used to identify a normal beat

if wave_R >0 && wave_R <= 26 && freq > 0 && freq <= 63.16 && intRR > 191 && intRR <= 374 **Type_Beat** = Normal Beat

4.1 Rhythm Classifier

The rhythm classifier makes groups of two to four beats and classifies the different rhythms. For this purpose the classifier uses a set of rules that reflect the properties of rhythms described in specialised literature [1]. We have extracted these rules, codified them in a programming language and tested them. One of those rules is shown in table 4. The parameters of the right (starting with "pf") are obtained from the users' data.

Table 4 represents the normal sinusual rhythm. It is the only rhythm that the system does not notify to the hospital because it is a desired rhythm. The meaning of this rule is the following: if the current beat and the previous one have their frequencies and PR intervals between the "pf" minimum and "pf" maximum and the R wave is smaller than the "pf" R maximum, then an only then is the rhythm normal sinusual rhythm. All these "pf" maximum and "pf" minimum values have been previously established by a physician.

Table 4. Rule used to identify a normal sinus rhythm

If (frequency_curr > pf_min_frec_normal && frequency_curr < pf_max_frec_normal && interval_PR_curr > pf_min_intervalPR_normal && interval_PR_curr < pf_max_intervalPR_normal && wave_R_curr < pf_max_R_normal && frequency_prev > pf_min_frec_normal && frequency_prev < pf_max_frec_normal && interval_PR_prev > pf_min_intervalPR_normal && wave_R_prev < pf_min_R_normal && interval_PR_prev < pf_max_intervalPR_normal) **Rhythm** = N (Normal Sinus Rhythm)

The classified rhythms correspond to the annotations found in PhysioNet (see table 5). All the rhythms are arrhythmias except normal rhythm N.

Table 5. Table of rhythm type

N	Normal sinus rhythm	VFL	Ventricular flutter
PREX	Pre-excitation (WPW)	AB	Atrial bigeminy
SBR	Sinus bradycardia	VT	Ventricular tachycardia
NOD	Nodal (A-V junctional) rhythm	B	Ventricular bigeminy
P	Paced rhythm	T	Ventricular trigeminy
IVR	Idioventricular rhythm	AFL	Atrial flutter
AFIB	Atrial fibrillation	BII	II heart block
SVTA	Supraventricular tachyarrhythmia		

Finally, beats and rhythms need to be stored by a Database and Communication Module. We explain this module in the following section.

5 Database and Communication Module

The database and communication module is the part of MOLEC that is in charge of storing the information generated by the Data Pre-processing and Decision Support modules and offers a set of services to the hospital. In our system, the PDAs work autonomously, receiving data from sensors, recording the data and analysing them in order to find anomalies. The PDAs are also able to send those anomalies, to receive queries formulated at the hospital about the signals, and to send the corresponding answers to the hospital.

The main goals of this module are: 1) to efficiently manage the restricted memory resources available in the PDA, at least when compared to the great capacity of ECG sensors to generate data ; 2) to reduce the quantity of data to be transmitted between the PDA and the hospital, taking into account that the wireless communication link used is usually more expensive than a wired one; and 3) to allow physicians in the hospital to formulate the kind of queries that they usually ask about ECG signals.

Flat files have been traditionally used to store recorded physiological signals. These data storage formats are not very interesting if high level queries want to be formulated, e.g. to find the number of isolated beats of type X in the current signal, or to find the number of sequences of Y consecutive beats of type X. However, these flat files are very appropriate if the goal is to reduce the quantity of data to be stored and/or transmitted, because they can be easily reduced to half their size after applying compression techniques. But these kinds of compression techniques have not been applied to classical signal files because they do not use an indirect model to analyse ECG signals: there is no computational node that processes and analyses the signal between the sensors and the host. Taking into account that, in our case, we do have an indirect model, it is possible to take advantage of it and to apply compression techniques. Therefore, we have applied common techniques in database and file technology in order to reach the previous goals: we use a relational database management system to store the data and to answer the queries made by the system (3rd goal mentioned above); and we also apply compression techniques to store the data that are not going to be queried (1st and 2nd goals mentioned above). At this point, we assume that physicians only make queries about abnormal beats and rhythms.

In figure 9, we show the database and communication module that consists of two submodules: 1) The *Data Manager Module*, which stores the data of the system; 2) the *Sender*, which communicates the PDA and MOLEC Hospital and builds the message in HL7 format [9]. HL7 is the first health care data-interchange standard.

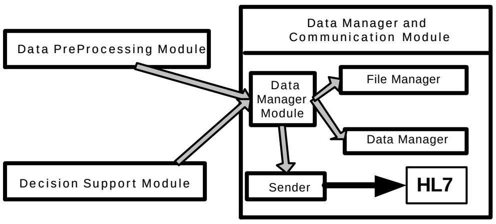

Fig. 9. Database and communication module

4.1. Data Manager

The Data Manager submodule receives the data from the Data Pre-processing module and Decision Support module and processes them according to the following criteria: all normal beats and rhythms are stored *compressed in files*; and all abnormal beats and rhythms are stored in a *relational database*. The following paragraphs explain these criteria, whose aim is to optimize the use of PDA memory and the use of the wireless network.

4.1.1. Process Followed by the Data Manager. The Data Pre-processing and the Decision Support Modules exchange messages though their interfaces with the Data Manager module, which interprets and manages these messages. The Data Manager receives a message every time a new beat is detected. Afterwards, it also receives the beat type and the current rhythm type, and decides how to store these data. To make this decision, the system implements the state transition diagram shown in figure 10.

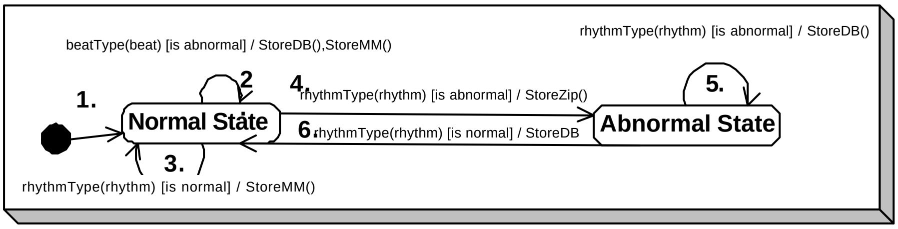

Fig. 10. State transition diagram

From left to right, figure 11 shows that: 1) the initial state is a normal rhythm; 2) if the system is in a normal state, but an abnormal beat (isolated beat) has been detected, it will be stored in the database (DB) and in a data structure in the main memory (MM). The system remains in a normal state; 3) if the system is in a normal state and a new normal rhythm has been detected, the state remains normal and the new data is stored in MM; 4) if the system is in a normal state, but an abnormal rhythm has been detected, the system retrieves from MM all data registered and creates a compressed file with them. The state of the system changes from normal to abnormal[3]; 5) if the system is in an abnormal state and a new abnormal rhythm is detected, the abnormal beat is stored in DB. The system remains in an abnormal state; 6) if the system is in an abnormal state, but a new normal rhythm is detected, the last abnormal rhythm is stored in DB, and the state of the system changes from abnormal to normal. The following subsection focuses on data compression. Afterwards, another subsection concentrates on the database and reports.

4.1.2. Compressed Data. As mentioned in this article, the digital signals emitted by the sensors are captured by MOLEC, in order to process them and determine possible abnormalities. Once the beats and the rhythms are analysed, it is possible to classify them in two great groups: *normal beats and rhythms* and *abnormal beats and rhythms.*

In figure 11, the proportion of normal and abnormal beats and rhythms corresponding to the recordings available in PhysioNet can be seen. Each one of the horizontal bars corresponds to a single patient: normal beats and rhythms are in grey, whereas abnormal ones in black. Abnormal rhythms and beats last an average of 6 minutes; some registers show no abnormal data (like register 100), and others are completely abnormal (like register 209). As for normal rhythms and beats, they last an average of 24 minutes.

In table 6, we can see the required memory size to store an ECG signal of some duration (Time column), depending on whether it is stored as: 1) a flat file used by PhysioNet and MIT-BIH (.DAT column); 2) the same compressed flat file (.DAT.ZIP column); 3) in a relational database (DB column). Notice that storing 6 minutes of abnormal signals in the database (1.6M) plus 24 minutes of normal signals compressed in a .dat.zip file (958K) is very similar to storing all 30 minutes in the .dat file (2.42M).

On the other hand, various ECG compression methods have already been developed. The most precise methods are less suitable because compression ratio is very poor, although the reconstruction has a high quality. Moreover, different compression methods yield different results, regarding that the compression ratio (2-30) and PRD (percent mean-square difference normalised by the original data) is just 0-30%. Hence, we only mention that we use ZIP compression, but a better one could be applied in our system if an appropriate tool for un/compressing were available.

Using our criteria, if the signal contains many *normal* beats and rhythms, the quantity of required memory is reduced to the half because zip compression is used; and if the signal contains many *abnormal* beats and rhythms, even more space could be required because that data would be stored at the PDA database (in order to be able to answer the kind of queries presented in the next subsection). Moreover, if more memory were needed at the PDA, part of the signal could be sent compressed to the hospital and uncompressed there. Physicians can query the PDA database about abnormal beats and rhythms from the hospital, or they can request to download in the hospital system the data from the PDA database.

[3] At this point, an alarm would be sent to the hospital system in order to indicate that a new abnormal rhythm has been identified.

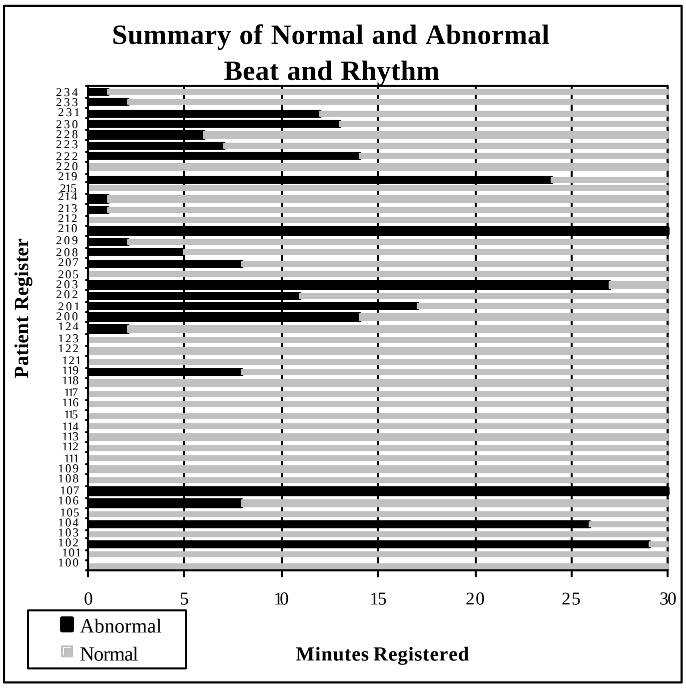

Fig. 11. Summary of recordings of MIT-BIH database

Table 6. Summary of the store types

Time	.DAT	.DAT.ZIP	DB	Time	.DAT	.DAT.ZIP	DB
2 sec	2.81 KB	1.55 KB	10 KB	6 Min	506.25K	236.64K	1.60 MB
10Min	843.75KB	398 KB	2.67 MB	24Min	1.98M	958.23K	6.42 MB
30 Min	2.42 M	1.17 M	8.3 MB				

4.1.3. Database Reports. Specialised literature [1] associates the functionality of holters with a set of reports, which enables physicians to analyse the data easily and quickly. In the following, we present some types of reports[4] usually offered by commercial products that show information about: 1) *automatic arrhythmia detection and identification* and 2) *ECG parameters*. In this subsection we present some SQL queries that would allow the system to generate the kind of reports previously mentioned. Notice that making these queries by using a database management system is much simpler than using flat files.

a) Automatic Arrhythmia Detection and Identification

These type of reports can be generated by writing queries that retrieve the different episodes of all abnormal rhythms, the frequency and the amount of beats involved. An

[4] There are other reports like analysis of ST segments evolution that we do not present due to lack of space.

example of an SQL query that obtains a summary of the cardiac rhythm levels (maximum, minimum, average) is shown in table 7. The columns indicate the rhythm type, the corresponding minimum and maximum rhythms, the average and the amount of beats involved. The last value is obtained by a "join" between the rhythm and beat tables. For this query, it would be interesting to define attribute RhythType as a clustering index, but that would require some reorganisation of the database.

Table 7. Levels of rhythm

```
SELECT     RhythmType as Type,  min(freq) as min,
           max(freq) as max, avg(freq) as avg, count(*)
FROM       Rhythm r, Beat b
WHERE      b.BeatNumber between r.firstBeatNumber
           and r.lastBeatNumber and  freq > 0
GROUP BY (RhythmType);
```

Type	min	max	avg	count	Type	min	max	avg	count
B	30.17	251.16	79.012	140	IVR	22.57	72.73	59.34	105
N	30.55	260.24	68.73	1467	SVTA	95.15	134.16	121.02	104
VFL	42.27	675.00	143.06	169	VT	74.48	183.05	103.97	4

b) ECG Parameters

Afterwards, we offer a detail of abnormal beats because the physician is interested in retrieving the number of sequences of consecutive n beats of type A, for n=1, 2, … m: number of isolated type A beats, number of couples of type A beats, and so on. This query is important because, this way, the physician can know the origin of the injured heart, depending on the beat type. The query can be repeated for supraventricular episodes (corresponding to beat types S, F, a and J) and ventricular episodes (corresponding to beat types V, ! and E).

In table 8, we present the SQL query and the obtained answer for the query corresponding to beat type A. The first column shows the length of the sequence, the second column is the number of sequences and the third one is the type of beat. Notice that this query consists of an autojoin in Beat table and of aggregation operators like "group by" and "count", that require some work. Hence, for this type of query it is interesting to define a clustering index of BeatNumber for Beat table. This is possible because consecutive numbers have been assigned to the beats.

Table 8. Detail of abnormal beats

```
SELECT   b.length, count(*), b.beatType
FROM     Beat b, Beat c, ECGRecording e
WHERE    b.beatType like '%A%' and c.beatType not like '%A%'
         and  c.BeatNumber = b.BeatNumber +1 and  c.ECGRecordID = e.id
         and b.ECGRecordID = e.id and  e.Name = 207 and
GROUP BY (b.length);
```

Length	count	Type	length	count	Type	length	count	Type
1	1	A	2	1	A	104	1	A

The previous queries can be formulated locally or remotely. If it is necessary to send information, the sender submodule sends a message to the hospital. This is explained in greater detail in the following subsection.

5.2 Sender

This module can send fragments of the user ECG signal to the hospital, as well as the reports solicited to MOLEC as seen in the previous section. A common standard is used to transmit medical data: HL7 [9]. An HL7 representation of clinical documents is called Clinical Document Architecture (CDA). The CDA is a document that specifies the syntax and semantics of a clinical document. It can include text, image, sounds and other multimedia content.

Table 9. Example of the message in HL7

```
MSH  | ^~& | SVL || SVC || 19900324101215 || ORU^W01| 19264|P|2.3 <cr>
PID  | 1 | 4567890 | 4567890 |     | Doe^John^Q^Jr^Mr <cr>
OBR  | 1 | 5678^SVC | 1234^SVL |    5^one-channel waveform recording^L <cr>
OBX  | 1 | CD|5&CHN^^L | 1| 1^ONE^0.5&mv^^200^-2048&2047 ||||||| F<cr>
OBX  | 2 | TS| 5&TIM^^L | 1| 19900324081237.525 F<cr>
OBX  | 3 | NA5&WAV^^L | 1| 0^1^2^3^4^5^6^7^8^7^5^4^3^2^1^0^1^F<cr>
OBX  | 4 | CE|5&ANO^^L | 1|    ^Beat Market ||||||| F || 19900324081237.525 <cr>
```

An HL7 message is composed of various segments (see table 9). The *Message Header* segment (MSH) describes the message, including the sender, the receiver, the subject and a message ID. The *Patient Identification* (PID) identifies the user whose data is being transferred and the *Observation/Result segment* (OBX) carries the data related to measured values. In the message of table 9, each waveform channel in a recording contains the channel definition (CHN), timing (TIM), the digital time series data (WAV) and the annotation (ANO).

When the message is received by the hospital, the physician reads the report and confirms the result obtained by MOLEC. Notice that there are tools that can show the data represented in HL7 messages to physicians.

6 Conclusions

Within the development of new technology that monitors people suffering from different illnesses, MOLEC is highly innovative because it provides the following advantages: 1) *Accessibility*: data recorded in MOLEC can be queried any time, locally or remotely. 2) *Promptness*: MOLEC detects anomalous rhythms, anywhere and anytime, as soon as they are produced, and sends the corresponding alarm to the hospital. Time can very often be vital in anomalous situations. 3) *Efficiency*: MOLEC optimises the use of wireless communications and PDA resources. 4) *Local analysis:* even if the wireless communication were unavailable, the signal could be analysed locally at the PDA. 5) *Openness*: MOLEC can deal with different kinds of sensors and PDAs, and it can cooperate with other systems that follow standards such as HL7 for clinical data communication. 6) *Simplicity*: making technical issues transparent to the users from the point of view of software and hardware components.

Acknowledgment

We want to thank Igor Beaumont, Lacramiora Dranca and Iñaki Inza for their help in the implementation of MOLEC and in the use of the classification tools.

References

1. Despopoulos, A.., Silbernagl, S.: Texto y Atlas de Fisiología. ISBN: 84-8174-040-3.
2. Daja, N., Relgin, I., Reljin B.,: Telemonitoring in Cardiology –ECG transmission by Mobile Phone. Annals of the Academy of Studenica 4, 2001.
3. Ventracor Limited. http://www.ventracor.co, 2003.
4. Cardio Control. http://www.cardiocontrol.com/cardio.htm, 2003
5. I. Sachpazidis. @Home: A modular telemedicine system. Mobile Computing in Medicine, Proceedings of the 2. Workshop on mobile computing. Heidelberg, Germany, April 2002.
6. Dimitri Konstansas Val Jones, Rainer Hersog. MobiHealth- innovative 2.5/3G mobile services and applications for healthcare. Workshop on Standardization in E-Health. Geneva, Italy. May 2003.
7. Kunze, C., Groβmann, U., Stork, W., Müller-Glaser, K.D.,: Application of Ubiquitous Computing in Personal Health Monitoring Systems. 36. annual meeting of the German Society for Biomedical Engineering, 2002.
8. MIT-DB Database Distribution. http://ecg.mit.edu/, 2003.
9. Health Level 7 (HL7): An application protocol for electronic data exchange in healthcare environments. http://www.hl7.org/ , 2003.
10. Rodríguez, J., Goñi A., Illarramendi, A.: Classifying ECGs in an On-Line Monitoring System. Submitted for Publication, 2003.
11. Physionet: “The research resource for complex physiologic signals“. http://www.physionet.org/, 2003.
12. CSE Database. France: http://www.iath.virginia.edu/~spw4s/CSE/CSE_database.frame.html, 2003
13. Jané, P., Blasi, A., García, J., Laguna, P.: Evaluation of an Automatic Threshold Based Detector of Waveform Limits in Holter ECG with the QT database”. Computers in Cardiology 1997, vol. 24, pp. 295-298.
14. Pan, J., Tompkin, W. J.: A real-time QRS detection algorithm". IEEE Trans. Biomed. Eng. BME-32: 230-236, 1985.
15. European ST-T Database: http://www.physionet.org/physiobank/database/edb/, 2003.