

# Transmitting critical Biomedical Signals over unreliable connexionless channels with good QoS using Advanced Signal Processing.

Stéphane Henrion\*, Corinne Mailhes\*, Francis Castanié\*

\*TéSa, Télécommunications Spatiales et Aéronautiques

2 rue Camichel, BP 7122 Toulouse Cedex 7, France

Tel (33) 05.61.58.80.08, Fax (33) 05.61.58.80.14.

Email : [stephane.henrion@tesa.prd.fr](mailto:stephane.henrion@tesa.prd.fr), Website : <http://ursafe.tesa.prd.fr>

**Abstract :** The U-R-Safe project (IST-2001-33352) aims at creating tele-medicine care environment for the elderly and convalescent. This objective is supported through the development of both a telecare concept and a technological mobile platform, for the home monitoring and health care provision. The idea is to provide a portable device which autonomously monitors different biomedical signals and is able to send and alarm to a medical centre if an abnormality is detected. In the initial version of U-R-Safe, only three sensors were included in the platform: ECG (electrocardiogram), SpO<sub>2</sub> and Fall Detector. Once an alarm is automatically set at the medical device(s) level (personal network) a transmission scenario is initiated. This one includes the transmission of some recorded data to a medical centre through a wide variety of networks and communication links (UWB radio-link, GPRS, or PSTN...). As no connectionless transmission is able to provide null error rate, when a deterministic QoS cannot be ensured as we use public networks, we studied the way to protect the transmitted biomedical data. In this paper, we will describe the philosophy and principle for transmission protection means and particularly some processing solutions for combined coding/restoration processes to recover useful information in case of data losses.

## 1. Introduction

Chronic conditions are becoming the first problem of public health in Western countries. Due to our current health care structures, too focused in the healing of acute conditions and provide little or no co-ordination among different health care levels, new emerging models of health care provision for chronic patients. To this end, the URSAFE system includes sensors (ECG, Oxygen saturation and fall detection) that monitor the patient condition. All are linked to a central unit, called Portable Base Station (PBS), through UWB (Ultra Wide Band) techniques. From the PBS data will be forwarded to a medical center over various Access Networks (terrestrial or satellite), according to their availability and coverage of the patient location.

When an alarm is detected at the monitoring devices level, a transmission scenario through the complete communication link is initiated. It includes the warning of a medical center through the transmission of few minutes of patient recorded biomedical data. Once those data received, a medical expertise may analyze the data, to confirm the diagnosis made at patient level and eventually activate an emergency situation.

Therefore the quality of the received data for the medical expertise is very important to provide an efficient solution.

As the transmission uses a wide variety of communications link, errors are liable to occur along the transmission chain. The way to ensure medically-readable information through protection (coding) means and associated restoration process is presented here.

## 2. Transmission context and issues

In the nominal configuration 3 devices are included in the URSAFE prototype : ECG, SpO<sub>2</sub> and Fall detector. The latter units provide mid term samples of a numerical which is compared to a medical threshold. But the ECG provides more complex and

complete information about the patient condition and shall be continuously analysed (over few minutes). Thus our scenario includes the transmission of the last minute of the recorded ECG to the medical centre for expertise.

According to the different characteristics of the transmission channels (bandwidth, memory storage capability) liable to be used between the sensor and the medical center, some additional signal processing can be required. These are mainly due to channel transmission bit rate capabilities (compression) and channel error models (coding and restoration) that could lead to partial data losses.

The different links ( fixed access network (PSTN), mobile access network (GSM/GPRS and future UMTS) or satellite interfacing (DVB-RCS technology)) included in this transmission chain are liable to induce errors on the transmitted data. These errors/losses can occur anytime and anywhere (according to the channel availability, memory overflows, protocol...) during a transmission process and can only be assessed in a statistical way. Indeed a fixed Quality of Service (QoS) does exist only if an end-to-end connection is reserved to the application. But this is not the case for the URSAFE project, which uses public networks between the local PBS and the medical center.

As no connectionless transmission is able to provide null error rate, when a quantified QoS cannot be ensured, the common way of obtaining the desired data when a transmission error is stated (through error detection coding<sup>3</sup>) is the use of a retransmission process. The receiver that noticed the transmission error asks for the re-sending of the data. But whatever the retransmission process (ARQ protocol) of the different layers of the protocols (either TCP in the IP one, or in 3 different sub-layers in GPRS), there is no way to provide null error rate (excepted if an infinite transmission delay is acceptable). According to the medical application of the URSAFE concept the maximum acceptable delay has been stated.

In its baseline form of the URSAFE concept the system is not dedicated to emergency cases (no information about the patient

location is included in the system). But in future developments of the system, more restriction about the transmission delay will be required for an efficient emergency service provision.

Thus we can sum up the transmissions chain (and its specificity) in the URSAFE concept concerning the ECG data (the longest to transmit) in two segments of 30 seconds each (before and after alarm event setting):

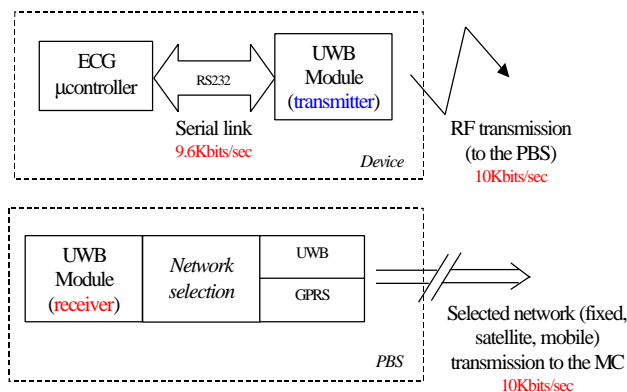


Fig.1

When a complete ECG message of 30sec length has been received at the PBS, the transmission to the medical center is initiated. One of the three possible networks is selected by the PBS. The data binary message length is organized in packets (which size can be parameterized between 0 – 2048 bytes) and transmitted on the selected network which uses a specific protocol.

If some errors (packet losses) are detected in the different layers of the data exchange protocol, then retransmission processes take place (based on the multiple retransmission ARQ protocol(s)), introducing delays in the global message transmission. It is not generally possible to intervene at the sub-layers level to modify the ARQ processes and thus to stop the multiple retransmissions. Thus according to the medical side (service needs) we will have to define a *time out point*, which is the maximum time delay until we consider the associated packet as lost. At Medical Center level, due to the use of an error detection code (CRC) on the transmitted binary data, we are able to know which packet is missing or lost and which packet is missing according to the *time out* delay.

Thus from the signal processing point of view, we can only base our approach on statistics of the missing packets, that is to say the packet loss probability. It may be possible to have this statistic as a function of the retransmission delay, but our inputs from our partners responsible for the transmissions over the mobile access: *“Packet loss does occur over GPRS links in both the downlink and up-link directions, but the incidence is relatively rare, and hard to quantify. When loss does occur, it is quite likely to occur in bursts of consecutive packets. As a general guideline, in core network design the packet loss probability is ideally set to 1% (however in real life can vary between 1-2% dependent on CDMA, FDMA, TDMA schemes and the factor mentioned above).”*

With such inputs we can study the compromise between the restoration algorithms at medical center level to reconstruct the missing data (lost packets) and the design of both the packet size and the data coding scheme.

### 3. Data organisation

Indeed if neither compression nor specific coding schemes were used, the loss of packets (which occur in bursts) will directly lead to a loss of data temporal samples in the received signal.

For instance, according to the provided statistics, with a packet length reduced to a single byte (*approximation*), the transmission of the 30 sec (9375 bytes) could result in the worst case in the loss of about 150 consecutive samples (in the worst case a single burst of 2% packets loss).

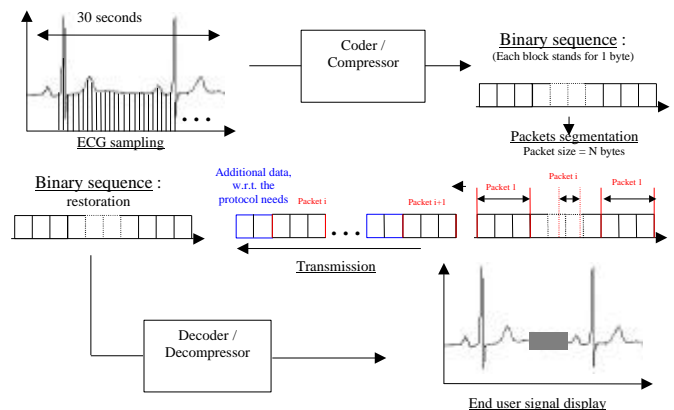


Fig.2

In another hand when the transmission is initiated, some additional bytes are added to the useful message (bytes for synchronization, identification, or CRC...). Therefore with a packet size reduced to a single byte (for instance if we consider 4 additional bytes for the protocol overhead) the total amount of data to be transmitted will be multiplied in a factor 7!

We can already point out the advantage of a specific coding which interleaves (or scrambles) the temporal data. This interleaving has to be done such that the successive packets don't include the successive temporal samples and the resulting burst(s) of packet losses result in smaller "holes" in the received and reconstructed temporal data. Thus, from the coding scheme point of view, it would be more efficient to reduce the packet size to the one of a single temporal sample. Then the resulting losses will result in independent holes (a majority of isolated samples) in the final (decoded) temporal data. But from transmission point of view the packet cutting up induces the transmission of additional data to the useful packet itself. Thus the reduction of the packet size to a single temporal sample degrades the global useful binary rate. Looking for this compromise with regards to the packet size is one of the objective of the coding / data organization study.

### 4. Strategy and solutions to overcome the transmissions losses

A defined QoS exists only if end-to-end connection is reserved to the application and this is not the case for URSAFE. Therefore, a strategy has to be decided. At the end user side, the medical doctor can receive data (ECG signals, mainly) with some possible "holes" in it, due to packet losses. The length of the missing samples is directly function of the lost packets length.

Two possibilities can occur:

- Either these packets arrive within an acceptable delay and the display is updated,
- Or these packets none not arrived after a maximum acceptable delay (acceptable for the medical end user, TBD).

In this last case, some restoration algorithms can be used in order to propose a way to restore the missing segments in the ECG display. Obviously, these restored samples will be highlighted (or displayed in a different color) in order for the medical doctor to be aware of the portions of ECG which have been restored. As soon as the missing packets arrive at the user end side, the restored samples will be replaced by the real data.

#### 4.1 State of the art of restoration scheme.

The reconstruction of partially known signals has been widely studied in the signal processing literature. A very good review can be found in [1]. Deeper studies can be found in [2] and [3].

In these papers, it is assumed that the signals are partially known in the sense that only a subset of its samples is supposed to be available for measurement. The problem is then to determine the signal from the available samples. However, the perfect recovery of missing samples from an arbitrary signal is of course impossible. This does not represent a difficulty since the signals we are faced with belong to the “real world” and often satisfy some known constraints. The constraint extensively used in recovery of missing samples is the frequency band-limitation. In the case of ECG signals, sampled at 250 Hz, this constraint can be considered. According to [4] it is clear that ECG signals can be assumed to be band-limited between 1 and 30 Hz.

In what follows, a signal with  $N$  samples is described by a  $N$ -dimensional vector  $x$ . The elements of the vector are denoted by  $x(0), x(1), \dots, x(N-1)$  and corresponds to the samples of the signal. The Discrete Fourier Transform (DFT) of the signal is the vector  $X$  defined by:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi \frac{nk}{N}}$$

A signal  $x$  is band-limited if its DFT vanishes on some fixed set, i.e.:

$$X(k) = 0 \quad \text{for } k \notin \{0, \pm 1, \pm 2, \dots, \pm M\}$$

Therefore, the band-limited signal is called a low-pass signal, band-limited to  $2M+1$  harmonics. In this case, it can be shown that:

$$x(n) = \sum_{k=0}^{N-1} X(k) \frac{\sin[\mathbf{p}(2M+1)(n-k)/N]}{N \sin[\mathbf{p}(n-k)/N]}$$

This relation is the base of most missing sample recovery algorithms. In order to represent the problem of missing samples, let us introduce a sampling operator  $D$  which maps a signal into another by setting to zero a certain subset of its samples. In matrix form, this corresponds to multiplication of the signal vector  $x$  by a diagonal matrix  $D$  containing only zeros or ones.

$$y = Dx$$

Given the observation  $y$ , how can we recover  $x$  from  $y$ ? In order that the problem be well-posed, a trivial necessary condition is that the number of samples of  $x$  that  $D$  preserves should be at least equal to  $2M+1$ , i.e. the dimension of the subspace of the low-pass signal  $x$ . The following proposed algorithms deals with that case.

For ECG signals,  $M$  is such that:

$$M = \frac{30}{250} N = 0.12N$$

Therefore, the above condition on ECG is that it remains at least 24% of samples. But note immediately that this represents a theoretical minimum condition. In practice, it is obvious (and also it can be demonstrated) that the efficiency of missing sample recovery algorithms will highly depend on how the missing samples are distributed over the total  $N$  samples. The worst case is when the missing samples are contiguous, as in the case of extrapolation problems. Scattered missing data usually leads to well-posed and stable reconstruction problems, whereas extrapolation leads to ill-posed, numerically difficult problems. In simulations and validations of the proposed algorithms, we will consider the worst case, i.e. the case where the missing samples are contiguous. But, it is important to note that the same reconstruction algorithms can be used in the case of scattered missing data with better results. Therefore, it will be interesting to propose a packet formation for ECG transmission in order to minimize the probability of missing samples to be contiguous (see section 3).

#### 4.2 The Papoulis-Gerchberg algorithm

The historical first method proposed for the missing data problem is the Papoulis-Gerchberg (P-G) algorithm. This is an iterative method, each iteration being based on a two-step procedure : filtering and re-sampling. The algorithm is initialized by

$$x^{(0)} = y$$

At each iteration  $k$ , the DFT of the vector  $x^{(k)}$  is computed and filtering is performed, band-limiting the DFT to  $2M+1$  non-zero harmonics. Then the inverse DFT restores the time domain knowledge in which we reinsert (by a sampling operation) the known samples of  $y$  and keep the estimated samples corresponding to the initial missing samples. This gives the new vector  $x^{(k+1)}$ . The iteration process can go on. An illustration of the two-step procedure (filtering by computing the DFT of the input signal vector and re-sampling by reinserting the known samples) is given below :

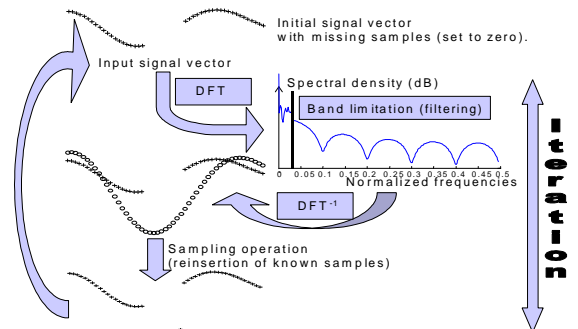


Fig.3

In order to validate this algorithm, it has been applied to a simulated signal: a filtered white noise such that  $M=16$  and  $N=512$ .

Simulations of missing samples have been achieved, assuming that the missing samples are contiguous. The main drawback of P-G algorithm lies in its convergence rate. When the missing sample number is small, a few iterations are necessary to get a quasi-perfect result.

However, the results of Papoulis-Gerchberg algorithm are less impressive when the number of missing samples is more important

The next figure presents the results obtained for 40 missing samples in the test signal. In this case, the number of iterations has to grow significantly in order to bring a satisfying result: after 1000 iterations, some error remains : the SNR is around 25 dB, inducing that this number of iterations should be higher:

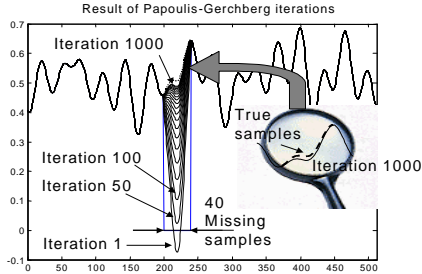


Fig.4

In order to fasten the convergence, we propose to initialize the Papoulis-Gerchberg algorithm *with a linear interpolation of missing samples* rather than with null values. This allows in the case of 40 missing samples, to end with a SNR of around 45 dB after 1000 iterations.

The major drawback of the Papoulis-Gerchberg iteration and similar methods lies in its convergence rate. The convergence has been established in the case of contiguous known samples. However, a more general algorithm has been proposed, linked to constrained iterative restoration, by introducing a relaxation factor,  $\mu$  and a sequence of successive approximations:

$$x^{(k+1)} = Tx^{(k)}$$

Where the operator  $T$  is defined by

$$T(.) = \mu y + (I - \mu D)B(.)$$

$y$  is the initial signal vector (with null values for missing samples),  $I$  is the identity matrix,  $D$  the sampling matrix (defined in the beginning of 2.1) and  $B$  is the operator corresponding to the filtering operation. Note that if  $\mu=1$ , this algorithm is equivalent to the classical Papoulis-Gerchberg one.

In order to optimize the convergence rate of the algorithm, the value of  $\mu$  can be computed depending on  $D$  and  $B$ .

This constrained iterative version of the Papoulis-Gerchberg algorithm has been also implemented. As waited, the convergence rate of the previous studied Papoulis-Gerchberg is improved.

At iteration 1000, the reconstruction is almost perfect, with a SNR of 60dB. Comparisons with Papoulis-Gerchberg algorithm are made on next figure, from a convergence rate point of view. Computing the optimal relaxation factor brings many improvements.

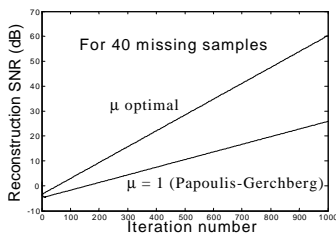


Fig.5

The previous iterative reconstruction algorithms are useful in connection with many reconstruction problems. However, whenever there are long contiguous gaps of missing samples, the convergence rate falls down to very low values. The minimum dimension formulations presented in next paragraph leads to better results

### 4.3 Minimum Dimension formulations

The computational requirements necessary to complete one iteration of the Papoulis-Gerchberg or its constrained variation is not dependent on the number of missing samples. These algorithms operate on  $N$ -dimensional vectors and the computational burden per iteration is mainly due to the filtering operation, which requires  $O(N \log_2 N)$  arithmetic operations using Fast Fourier Transform. The first interest of the minimum dimension method is that it leads to sets of linear equations for the unknown samples, with as many equations as unknown samples, thus less cost-computive than Papoulis-Gerchberg algorithms.

The basic formulation of minimum dimension restoration algorithms lies in the sampling theorem. Any band-limited signal with a spectral band  $Df$  (the spectral contents of the signal lies between  $-Df/2$  and  $+Df/2$ ) can be written as:

$$x(t) = \Delta f \sum_{n=-\infty}^{+\infty} x(n) \frac{\sin(\Delta f(t-n))}{\Delta f(t-n)}$$

Where  $t$  is an arbitrary time instant.

This well-known formula can be applied in the discrete case of an  $N$ -dimensional signal with unknown and known samples. Let us note  $U$  the subset of indices of unknown samples and  $K$  the subset of indices of known samples. The above formula can be split in two sums: one taking into account unknown samples, the other one on the known samples:

$$x(k) = \Delta f \sum_{n \in U} x(n) \text{sinc}(\Delta f(k-n)) + \Delta f \sum_{n \in K} x(n) \text{sinc}(\Delta f(k-n)) \text{ for } k \in U \text{ or } k \in K.$$

This formula is true on every sample  $x(k)$ , either known or unknown. But we focus our interest in using this equation for each unknown sample. This leads to a set of equations, the number of which is equal to the number of unknown samples. This set of equations can be written in a matrix form:

$$x_{miss} = Sx_{miss} + h$$

Where  $x_{miss}$  represents the missing sample vector,  $S$  is a matrix of cardinal sinus (a sub-matrix of the Papoulis-Gerchberg matrix  $T$ ) and  $h$  is a vector computed from known samples. Then, if  $I-S$  is non-singular, any solution for linear equation system can be used to solve this equation. Iterative and non-iterative algorithms exist.

We have implemented the minimum dimension algorithm in a non-iterative form. The results on the test signal are so good that it is impossible to represent a reconstruction SNR since "the reconstruction error remains almost null for a length of contiguous missing samples from 1 to 150!" For a length of contiguous missing samples higher than 150, some ill-posed problem appears: the matrix  $I-S$  is close to be singular.

Therefore, this kind of algorithms seems very interesting, even if it presents some unstability (but in extreme cases).

#### 4.4 Validation on ECG signal

Even if ECGs are band-limited signals, they are non-stationary which may degrade the proposed algorithm performances. Moreover, restoration algorithm results may highly depend on the *position* of the missing samples. Therefore, all proposed algorithms have been tested on portion of 512 points of ECG signals and for different localizations of the missing samples.

Firstly, Papoulis-Gerchberg algorithm (P-G algo.) has been applied. The next figure presents some results on restoration error power versus length of missing samples, for some tested localizations of the missing samples and for different iteration number.

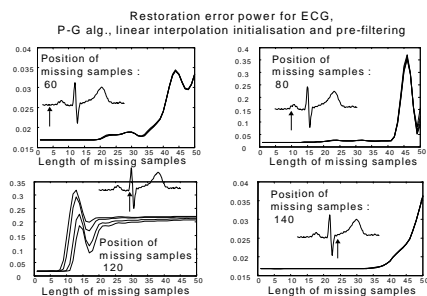


Fig.6

On each sub-figure, 4 curves have been drawn, depending on the iteration number of the P-G algo (10, 100, 500, 1000). For some sub-figures, these 4 curves are superimposed, meaning that the number of iterations has not to be very high in order to get good results. In the the worst case, the number of iterations is of importance and the maximum length of missing samples that can be recovered is around 5 to 10, depending on the iteration number.

This represents a result of maximum 1% of 2% of contiguous missing samples which can be recovered using the Papoulis-Gerchberg algorithm. Note that if the missing samples where not contiguous, obviously, the percentage of missing samples which can be recovered by this algorithm would be higher.

The constrained iterative restoration algorithm which belongs to the same class of algorithms than the Papoulis-Gerchberg one gives similar results on ECG signals. Note that the computing cost of optimum relaxation factor may be a strong drawback of this algorithm, since the restoration algorithms are used while the user end system will wait for the lost packet to arrive.

The last proposed algorithm, the minimum dimension formulation, performs very well on ECG signals, using the proposed modification (pre-filtering with an ideal filter at the assumed cut frequency, here, 30 Hz). The problems of ill-conditioned matrix appear for a number of contiguous missing samples exceeding 45 (over 512). This seems a value above which the transmission problem should not be. However, if error propagation studies lead to consider that such percentage of contiguous missing samples can have a non-negligible probability, solutions exist, by implementing sophisticated algorithms for linear system solving.

What is more important, is that the results given by this algorithm are far better than the one obtained with Papoulis-Gerchberg based

one. First, the evolution of the restoration error power as a function of the number of missing samples does not depend on the position of these missing samples, as was observed for Papoulis-Gerchberg algorithm. Whatever the position, the error has the same shape than the one presented below. We have chosen to present the results obtained when the missing samples are located on the QRS wave. As can be seen, the error power is less than with the P-G algorithm and allows considering that *a maximum of 40 contiguous missing samples can be recovered (over 512)*.

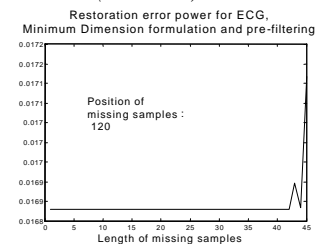


Fig.7

#### 4.5 Data organisation to be compliant with the restoration capabilities

From the previous results, at user end level, the restoration process is able to restore 40 successive missing samples within a window width of 256 samples. As the data are basically coded on 10bits, this allows a loss of 400 bits or 50 bytes.

Therefore the maximum data message length to transmit in the URSAFE project (without compression) is 75Kbits or 9.375Kbytes (30 sec of recorded ECG). Thus the maximum data length liable to be lost is around 200 bytes !

Consequently it necessary to consider a specific data organization prior to transmission. For instance, such a code could start with the direct time binary coding followed by a packet segmentation with a length equal at most to 50bytes (40 temporal samples). We could then permute these packets to decorrelate their organization. Then if losses occur, in the case of a **burst** of 187bytes losses (nearly 4 packets) the resulting temporal sample losses won't be contiguous. Indeed after decoding packet losses will result in "reduced" holes of 40 samples (50bytes) but not contiguous. To be more protective we can think of margin on the packet size, if we reduce it to 40bytes (32 temporal samples)...

The resulting specifications for the data-coding scheme are the following:

- packet length : less than 50 bytes
- coding : shall ensure that 2 successive temporal packets are more distant than 4 (number of successive packet losses)\*N=187bytes/packet\_length packets in the resulting code.

Thus, we propose to divide the total message length (30s of ECG) in blocks of K temporal samples. Then an interleaving code of these blocks is performed before binary packing and transmission (M blocks in a packet). According to the maximum data length restoration capability, (40 from section), the block sizing can be considered between 1 and 40 samples.

The propagation of packet losses through the decoded data is illustrated on the Fig.8 (In this figure, it is assumed that only one packet is corrupted (in red) during the transmission):

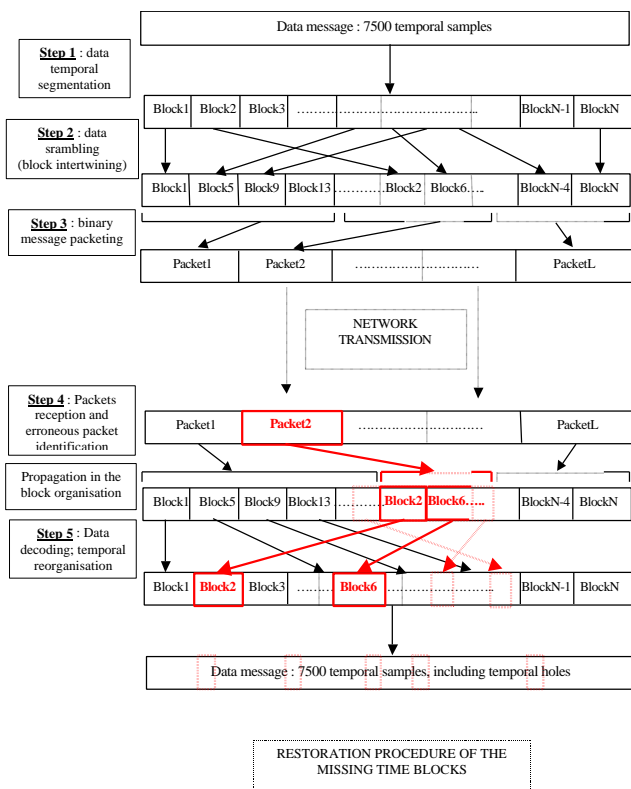


Fig.8

The aim of the interleaving is to ensure that the resulting “holes” (in the recovered signal) have a length below 40 samples and are separated at least by 256 samples, according to the restoration capabilities. For that purpose, we have studied the compromise on block length, packet size and spacing between two consecutive blocks in the interleaving coded message. For our application, one good compromise seems to be the following interleaving scheme :

- a block length of  $K=20$  samples (less than max restorable hole length=40)
- $M=25$  intertwined blocks leading to 15 blocks sequences along the whole 30 seconds of ECG
- a packet length of 256 bytes (according to the UWB link).

With such a coding scheme, simulations results showed that even for 8% of packet losses the combined decoding/restoration scheme allow to reconstruct a more functional signal for a medical expertise. On Fig.9 the first signal is the one including burst packet losses effect without interleaving coding (signal not available = holes); the second shows the same amount of losses effect obtained with the interleaving scheme and the latter the reconstructed signal (reconstructed parts are highlighted in red). A specific zoom on original signal, reduced losses and reconstructed part is also proposed.

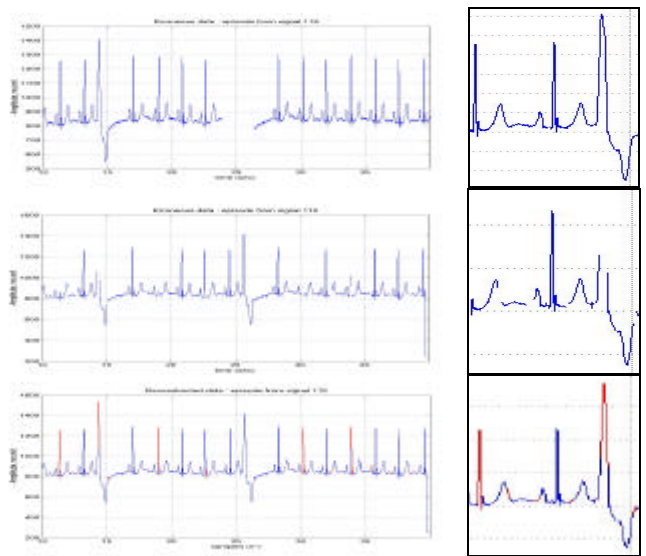


Fig.9

The efficiency of the proposed restoration scheme was proposed to medical expertise who recognized the efficiency of the process if the reconstructed parts were specifically labelled.

## 5. Conclusions

All the proposed SW solutions have been presented and evaluated through simulation results with signals adapted to our application objectives : transmission of 30 sec of recorded ECG.

We defined an efficient solution to answer non “in mind”, but effective, packet losses issues among the access networks.

As the global restoration process is highly modular, since it can be easily apply to bandpass limited signal and not only to ECGs. Moreover or combined interleaving/restoration principle can be designed and parameterized (in terms of blocks length, packet size) to fulfill other transmission channel capabilities in terms of *loss probabilities*.

### References :

- [1] : Marvasti, Nonuniform Sampling : theory and practice, *Kluwer Academic/Plenum Publishers*, New York, 2001.
- [2] : P.J.S.G. Ferreira, Interpolation and the Discrete Papoulis-Gerchberg Algorithm, *IEEE Trans on Signal Processing*, vol 42, n 10, October. 94, pp 2596-2606.
- [3] : P.J.S.G. Ferreira, Interpolation in the time and frequency domains, *IEEE Signal Processing Letters*, vol 3, n 6, June 96, pp 176-178.
- [4] : Laurent Clavier, “Analyse du signal electrocardiographique en vue du dépistage de la fibrillation auriculaire”, *A thesis*