Prediction of time series using ARMA models in an energy-efficient body area network

Karel Heurtefeux
Qatar Mobility
Innovations Center
Doha, Qatar
Email: karelh@qmic.com

Nasreen Mohsin
Qatar Mobility
Innovations Center
Doha, Qatar
Email: nasreenm@qmic.com

Hamid Menouar
Qatar Mobility
Innovations Center
Doha, Qatar
Email: hamidm@qmic.com

Najah AbuAli
College of Information Technology
UAE University,
Al Ain, United Arab Emirates
Email: najah@uae.ac.ae

Abstract—This paper investigates the tradeoff between accuracy and complexity cost to predict electrocardiogram values using auto-regressive moving average (ARMA) models in a fully functional body area network (BAN) platform. The proposed BAN platform captures, processes, and wirelessly transmits six-degrees-of-freedom inertial and electrocardiogram data in a wearable, non-invasive form factor. To reduce the number of packets sent, ARMA models are used to predict electrocardiogram (ECG) values. However, in the context of wearable devices, where the computing and memory capabilities are limited, the prediction model should be both accurate and lightweight. To this end, the goodness of the ARMA parameters is quantified considering ECG signal, we compute Akaike Information Criterion (AIC) on more than 900000 ECG measures. Finally, a tradeoff is given accordingly to the hardware constraints.

Keywords—Body area network, autoregressive moving average, akaike, energy efficiency

I. INTRODUCTION

Wireless Body Area Networks (BAN) are formed by several low-energy wirelessly interconnected biomedical or inertial sensor devices. The sensors usually capture various physiological parameters of the human body (*e.g.* temperature, heart rate, Electroencephalography (EEG), Electrocardiography (ECG), blood pressure, blood oxygen saturation, etc.). It also determines parameters of the physical environment, such as the amount of sunlight exposure or ambient air quality.

In a typical BAN architecture, sensor data are transmitted wirelessly to a coordinator (also called aggregator or gateway) where the data are forwarded to an access point and then, sent over Internet to a remote medical server for storage and analysis. Due to constraints such as energy and computation capability, non-deterministic sensor failures, radio links instability, and distrusted environments, designing and deploying a robust BAN platform are some of the challenges encountered.

Motivation. BANs have become a leading approach for several promising applications in the medical and healthcare fields. But despite the rich availability of research works, there are only few fully functional applications that can be actually deployed in real scenarios. In particular, limited resources in energy and in radio communications make real-world deployment difficult.

Contributions. This article is based on a preliminary work [1] investigating the performances of the proposed architecture in terms of accuracy and efficiently, based on three different thresholds. The filter proposed in [1] is based on similarity of two consecutive packets. The similarity filter has shown to be efficient in reducing traffic of movement signals but remains inoperative for ECG signals due to their dynamic nature. In this article, we propose to investigate a more complex filter using auto-regressive moving average (ARMA) models in order to limit the number of packet sent by the ECG sensor.

Novelty. WBANs have recently been the subject of intense research by many researchers worldwide and interesting research results have been already produced in that topic and especially in the areas of physical and network layer mechanisms. However, there are only few studies related to the development of practical, efficient and low-energy WBANs system. ARMA models have been used in several domains such as financial analysis, meteorology or urban Traffic. However, only few studies exist to predict ECG. For instance, in [2], the authors base their prediction on simulated ECG signals. In this paper, ARMA model is used and evaluated in terms of AIC and RMSE. The proposed evaluation is based on more than 900000 real measurements obtained using Shimmer 4-leads ECG sensors [3]. In order to use ARMA model for ECG forecasting, appropriate model parameters have to be selected. ARMA models provide a description of a stochastic process in terms of two polynomials, one for the auto-regression and the second for the moving average. The $ARMA(p, q)$ takes two parameters: p is the order of the auto-regressive part and q is the order of the moving average part. To select the appropriate order p and q and quantify the goodness of fit on the ECG measures, the Akaike Information Criteria (AIC) [4] is used. AIC measures of the relative quality of a statistical model for a given set of data.

II. MODELS

A. Auto-Regressive Moving Average Models

The process $\{x_t\}$ is an auto-regressive moving average process of order p, q if there exist constants $a_1, ..., a_p, b_1, ..., b_q$ such that:

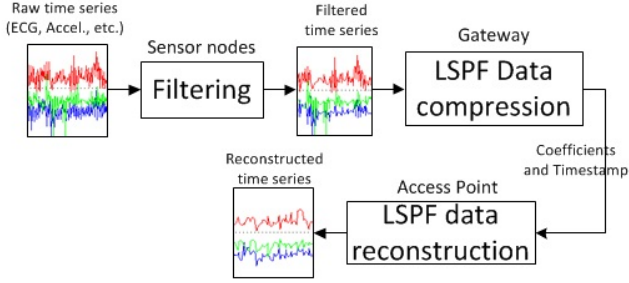$$x_t = \sum_{j=1}^{p} a_j x_{t-j} + \sum_{j=1}^{q} b_j \varepsilon_{t-j} + \varepsilon_t \qquad (1)$$
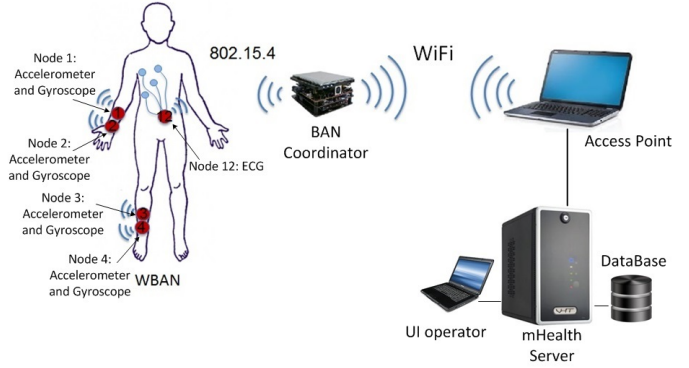
Fig. 1: Flowchart of the BAN architecture

Fig. 2: Overview of the BAN Architecture

where $\{\varepsilon_t\}$ is zero-mean white noise, and ε_t is uncorrelated with $x_{t-1}, x_{t-2}, ...$. The $ARMA(p,q)$ process exists and is weakly stationary if and only if all the roots of the polynomial $P(z) = 1 - a_z - ... - a_p z^p$ are outside the unit circle. The process is said to be invertible if all the roots of the polynomial $Q(z) = 1 + b_1 z + ... + b_q z^q$ lie outside the unit circle. A time series is invertible if and only if it has an infinite-order autoregressive $AR(\infty)$ representation of the form:

$$x_t = \sum_{j=1}^{\infty} \pi_j x_{t-j} + \varepsilon_t \tag{2}$$

where π_j are constants with $\sum \pi_j^2 < \infty$.

In the linear prediction problem, we want to forecast x_{t+h} based on a linear combination of $x_t, x_{t1}, ...,$ where $h > 0$ is the lead time. It can be shown that the best linear forecast, *i.e.* the one which makes the mean squared error of prediction as small as possible, denoted by $\hat{x}_{t+h}$, is determined by two characteristics: (1) $\hat{x}_{t+h}$ can be expressed as a linear combination of $x_t, x_{t1}, ...$ and (2) the prediction error $x_{t+h} \hat{x}_{t+h}$ is uncorrelated with all linear combinations of $x_t, x_{t1}, ...$

B. Akaike Information Criterion

In WBANs, where devices are strongly constrained in terms of computational power, it is important to limit the complexity of the model. The Akaike information criterion (AIC) measures the relative quality of a statistical model for a given set of data.

$$AIC = 2k - 2\ln(L) \tag{3}$$

where k is the number of parameters in the model, and L is the maximized value of the likelihood function for the model (*i.e.* how likely a particular population is to produce an observed sample). In short, AIC deals with the trade-off between the accuracy of the model and its complexity.

III. OVERVIEW OF THE PROPOSED PLATFORM

This section describes the wearable BAN platform which was designed, implemented and evaluated to enable the remote monitoring of workers in harsh environments. The target application scenario is first described below, followed by an overview of the hardware, software and communication components.

A. Application Scenario, communication architecture and materials

In this study, we focus on the remote monitoring of workers in harsh environment. With the expansion and emergence of large mega construction projects, the safety and health of workers is becoming a serious concern worldwide. For instance, the number of deaths due to work-related accidents or diseases, remains unacceptably high at around 2.3 million per year [5].

In this context, it is expected that WBAN technology will enhance the safety and health of workers, for example, by enabling the remote monitoring of workers in unhealthy environments [6]–[9]. As a result, there is a need to monitor physiological signs (*e.g.* body temperature, pulse rate, respiration rate, blood pressure, etc.). With wearable sensors and BANs, workers can be monitored remotely and quick assistance can be given if anomalies on the vital signs are detected. To be practical in such context, a BAN should be able to send data continuously to a remote server for storage and analysis, while being energy efficient and accurate. In that purpose, low-power technologies, compression techniques and filtering are targeted. Moreover, the monitoring of the body movements (e.g. acceleration, orientation, etc.) can be useful in implementing safety related algorithms, such as fall detection or activity recognition, and thus ensuring the safety of the workers.

In this paper, wearable sensors communicate using the IEEE 802.15.4 standard, which covers the PHY and the MAC layers. Due to its good performance in terms of energy consumption, such a standard is a good candidate for constrained devices such as battery-powered wearable sensors and it is the basis for numerous specifications such as ZigBee [10], WirelessHART [11], or ISA100.11a [12]. In the proposed platform, a BAN coordinator aggregates the traffic coming from the sensor nodes and forwards it to an access point. IEEE 802.15.4 is used for the on-body communications between the sensors and the BAN coordinator, and IEEE 802.11 / WiFi is used for the communication between the BAN coordinator and external access points, as shown in Figure 2. In IEEE 802.15.4, the CSMA/CA MAC protocol is generally used by

sensors to send data and they can theoretically transmit up to 250 kbs at 2.4 GHz which is a sufficient data rate for typical wireless sensor applications. However, this MAC layer implements a collision avoidance mechanism based on random backoff which is not efficient for periodic and real-time traffic and creates latency and collisions. As a result, we propose a new MAC layer described in detail in the following section.

In order to enable efficient on-body communications in terms of latency, delivery ratio and energy consumption, we designed and implemented two specific algorithms: a filtering algorithm running on the WBAN sensor device, and a data compression algorithm running on the coordinator.

The system used for the testing and validation of our solutions consists of three components:

Sensor Nodes consists of five Shimmer nodes [3]. The Shimmer node is a small sensor platform well suited for wearable applications. It has low-power communication capabilities enabling long-term data acquisition and real-time monitoring. In this work, four nodes integrate 3-axis accelerometer and 3-axis gyroscope and one node is dedicated to heart monitoring and integrates a 3-lead ECG. Each node runs on TinyOS [13]. The characteristics of Shimmer nodes are summarized in Table I.

Coordinator Node consists of: (i) a Beagleboard XM [14], (ii) a BeagleTouch Screen (iii) an 802.11 module for Wi-Fi connection and (iv) an 802.15.4 module for Zigbee connection. The Ubuntu 11.10 OS is used to run the coordinator. A lightweight server is implemented on the platform to perform the forwarding and the polynomial data compression (LSPF). The coordinator characteristics are summarized in Table I.

Access Point which carries the proper storage, database and application software. It is intended to be constantly available (*i.e.* 24/7) and scalable to enable the monitoring of a large number of patients. The server runs real-time analysis of sensor's data, provides user access to the database at various levels (e.g. patients, relatives, physicians, etc.) and generates alarm in case of emergencies.

The signal is first captured, amplified and digitized by Shimmer nodes. It is then quantized at the selected sampling frequency ranging from $1Hz$ to $1kHz$. Next, the shimmer node transmits the data (i.e. 7 samples per packet) to the coordinator which forwards the packet to the access point.

B. Filtering Algorithm

As illustrated in Figure 2, sensor nodes gather sensor information and communicate with the BAN coordinator. With respect to their constraints in computational power, a lightweight filtering algorithm is implemented. It is defined to limit the amount of data sent by the sensor nodes. To form a packet, each sensor aggregates 5 values of each signal (e.g. acceleration and angular velocity for 6-axis sensors, lead I and lead II for the ECG sensor). Then, it sends the packet to the coordinator. The filtering algorithm proposed in this study only considers ECG signal. It computes the prediction $x_{t+1}, ..., x_{t+5}$ based on $ARMA(p, q)$. Then, the predicted

Device Manufactor	Shimmer Node	Beagleboard-Xm
Microcontroller	MSP430	AM37x 1GHz ARM Cortex-A8
Radio	TI CC2420 (802.15.4) [15] and RN Bluetooth module	TI CC2420 (802.15.4) and Ralink RT2571WT (802.11.b/g)
TX Power	0dBm	802.15.4: 0dBm, Wifi: 13dBm
Radio sensitivity	-95dBm	802.15.4: -95dBm, Wifi: -70dBm
TX/RX consumption	17.4mA/18.8mA	802.15.4: 17.4mA/18.8mA, Wifi: 390mA/270mA
Battery	280mAh, 3.7v	8400mAh, 5V
Sensing capabilities	3-axis Accelerometer, 3-axis Gyroscope, ECG	None
OS	TinyOS	Ubuntu 11.10
MAC protocol	CSMA/CA and Dynamic TDMA	CSMA/CA (WiFi and 802.15.4) and Dynamic TDMA (802.15.4)
Protocol thresholds	Quadratic distance threshold 2%	RMSE Threshold: 1%, Maximum polynom order: 10

TABLE I: Summary of the platform characteristics.

values are compared to the measured values. The sensor defines the similarity by computing the Root Mean Square Error (RMSE) between them. Let two vectors v and w in R^n be as follow: $v = (v_1, v_2, ..., v_n)$, $w = (w_1, w_2, ..., w_n)$. The RMSE is: $RMSE(v, w) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (v_i - w_i)^2}$. If the RMSE is higher than a given threshold, the measured packet is sent to the BAN coordinator. At the BAN coordinator level, the prediction is computed based on the previous received packets. If a new packet is received the predicted values are replaced by the measured ones.

IV. EXPERIMENTAL RESULTS

A. Methodology

The results have been obtained by mathematical computation using Matlab software and based on real-world data. The data which have been obtained with an ECG Shimmer sensor contains more than 900000 samples (i.e. 90000 packets). The orders p and q varied between 10 (i.e. 2 packets) and 50 (i.e. 10 packets).

B. ARMA performance

AIC. Fig.3 illustrates the Akaike Information Criterion (AIC) according to the orders of auto-regressive p and moving average q. The higher AIC values (*i.e.* the worst trade-offs) are represented in red, while the lowest AIC values (*i.e.* the best trade-offs) are represented in blue. Practically, p and q represent the number of packets that should be stored in the node in order to compute the model. As a result, p and q represent the overhead in terms of computational power (model complexity) and memory usage (as previous packets should be stored). The best accuracy-complexity trade-offs is obtained with $p = 40$ and $q = 35$ (*i.e.* 8 and 7 packets respectively) as shown on Fig.3. Note that the gain obtained in terms of AIC is significant for the orders under $p = 15$ et $q = 15$; however, beyond these values the gain decreases.

Mean Error. Now, we illustrate the evolution of the mean error in detail according to the orders p and q to determine their optimal values. To do so, we compute the difference between
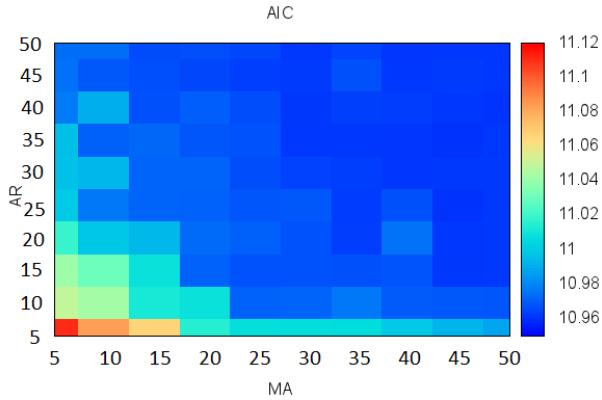
Fig. 3: Akaike Information Criterion according to Auto-Regressive and Moving Average orders
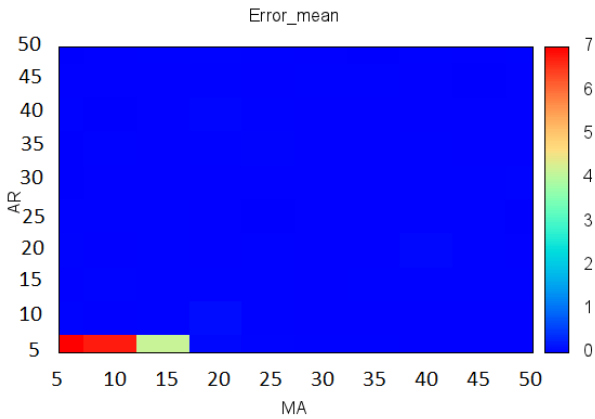

Fig. 4: Average Error of prediction according to Auto-Regressive and Moving Average orders

the values obtained through model $ARMA(p,q)$ and those obtained experimentally. As shown in Fig. 4, the mean error is between 0 and 1 for the values $p > 5$ and $q > 15$; however, beyond these values the gain is small.

Due to the constraint capabilities of the network devices, the model complexity should be minimized in order to decrease the memory and computational power usage. Here, the model ARMA(2,3) seems appropriate both in terms of performance and complexity.

V. CONCLUSION AND FUTURE WORK

In this paper, a study was conducted on ECG signal prediction based on ARMA model. We established orders p and q to reach a tradeoff between model complexity and accuracy by computing AIC and the average prediction error given based on more than 900000 ECG measures. The relatively low orders found leads us to think that an implementation in a constrained embedded systems is feasible.

Future works. The current study does not investigate implementation of ARMA-based prediction on Shimmer nodes.

It certainly would be interesting to propose such algorithm dealing with constraints in terms if computational power and memory space. In addition, the efficiency of the prediction to reduce the number of transmitted packets should be evaluated regarding its cost in terms of energy consumption.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] K. Heurtefeux, E. Ben Hamida, and H. Menouar, "Design and implementation of a sustainable wireless ban platform for remote monitoring of workers health care in harsh environments," in *New Technologies, Mobility and Security (NTMS), 2014 6th International Conference on.* IEEE, 2014, pp. 1–5.

[2] J. J. Aguila, E. Arias, M. M. Artigao, and J. J. Miralles, "A prediction of electrocardiography signals by combining arma model with nonlinear analysis methods," in *Proceedings of the 10th WSEAS International Conference on Applied Computer and Applied Computational Science*, ser. ACACOS'11. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2011, pp. 31–37. [Online]. Available: http://dl.acm.org/citation.cfm?id=1965610.1965614

[3] Available online: www.shimmersensing.com, January 2014.

[4] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on automatic control*, vol. 19, pp. 716–723, 1974.

[5] WHO, "Health indicators of sustainable jobs," 2012.

[6] L. Senyurek, K. Hocaoglu, B. Sezer, and O. Urhan, "Monitoring workers through wearable transceivers for improving work safety," in *Intelligent Signal Processing (WISP), 2011 IEEE 7th International Symposium on*, Sept 2011, pp. 1–3.

[7] A. Ceccarelli, A. Bondavalli, J. Figueiras, B. Malinowsky, J. Wakula, F. Brancati, C. Dambra, and A. Seminatore, "Design and implementation of real-time wearable devices for a safety-critical track warning system," in *High-Assurance Systems Engineering (HASE), 2012 IEEE 14th International Symposium on*, Oct 2012, pp. 147–154.

[8] J. Forsyth, T. Martin, D. Young-Corbett, and E. Dorsa, "Feasibility of intelligent monitoring of construction workers for carbon monoxide poisoning," *Automation Science and Engineering, IEEE Transactions on*, vol. 9, no. 3, pp. 505–515, July 2012.

[9] S. Li and B. Gerber, *Evaluating Physiological Load of Workers with Wearable Sensors*, ch. 51, pp. 405–412. [Online]. Available: http://ascelibrary.org/doi/abs/10.1061/9780784412343.0051

[10] Z. Alliance, "Zigbee specification," Tech. Rep., Jun. 2005.

[11] http://www.hartcomm.org.

[12] https://www.isa.org/isa100.

[13] Available online: www.tinyos.net, January 2013.

[14] Available online: http://beagleboard.org/hardware-xm, December 2013.

[15] "Chipcon inc, cc2420 datasheet," Available online: http://www.ti.com/product/cc2420, February 2008.