

RA-SAX: Resource-Aware Symbolic Aggregate approXimation for Mobile ECG Analysis

Hossein Tayebi, Shonali Krishnaswamy, Agustinus Borgy Waluyo, Abhijat Sinha
 Authors
 Centre for Distributed Systems and Software Engineering
 Monash University
 Melbourne, Australia
htayebi@gmail.com,
Shonali.Krishnaswamy@monash.edu, {abwaluyo,
callmerahul20}@gmail.com

Abstract—There is a growing focus on 24/7 cardiac monitoring that leverages state of the art mobile phones and commercial-off-the-shelf (COTS) wearable bio-sensors. While many signal processing techniques for mobile ECG analysis have been developed, these techniques tend to be computationally intensive. In this paper, we propose, develop and evaluate a resource-aware and energy-efficient time series analysis technique for real-time ECG analysis on mobile devices based on the well-known SAX (Symbolic Aggregate Approximation) representation for time series termed RA-SAX.

Keywords: Mobile Devices, ECG, Time Series

I. INTRODUCTION

In recent years, the number of people who suffer from cardiovascular diseases has significantly increased. With the latest advances in technology, new ways of improving a patient's well-being have become possible. A major area of work which raises a wide range of possibilities is real-time monitoring systems using bio-sensors and mobile devices like PDAs/Smart Phones [1]. This can be crucial to out-of-hospital's patient care.

Traditional approaches to cardiac monitoring involved treating the mobile device primarily as a transmission/communication device which was collecting and sending the signals received from the bio-sensor to a remote server for analysis [1]. However, with the advances and increasing sophistication of mobile devices, coupled with the cost/battery drain of continuous data transmission from mobile devices, there is an increasing focus on performing the signal processing of the ECG signals from the bio-sensors using the mobile phone.

In this paper, we propose the use of Symbolic Aggregate Approximation (SAX) [2] which is known to be a computationally efficient time-series representation and has been used for cardiac analysis (albeit in a non-mobile setting) for ECG analysis in a mobile environment. Studies such as [3] have shown that the key to effective mobile data mining on mobile devices is to perform “resource-aware adaptation”. This paper presents our resource-aware adaptation of SAX.

Mohamed Medhat Gaber
 School of Computing, University of Portsmouth
 Portsmouth, United Kingdom

Mohamed.Gaber@port.ac.uk

We experimentally demonstrate the impact of the resource-adaptation of SAX on the battery life of a mobile device. We then show the use of RA-SAX for mobile ECG analysis by presenting evaluations of RA-SAX for classification, analysis using K-Nearest Neighbor (K-NN) and an enhancement using a combination of lightweight clustering and KNN. We present the experimental evaluation as resource-levels change in a mobile setting.

The paper is organized as follows. Section 2 is dedicated to describing our Resource-Aware SAX-based framework for mobile ECG analysis. Section 3 discusses the implementation and presents the experimental evaluation. Section 4 concludes the paper.

II. RESOURCE-AWARE ECG ANALYSIS FRAMEWORK

Our resource-aware ECG monitoring system is designed to provide local classification on the mobile devices such as a smart phone. It consists of three main components (Figure 1): (i) Data Acquisition Module which is responsible for acquiring stream of ECG information collected from a biosensor that records live ECG signals of a patient and transmits them to a handheld device for further analysis (ii) RA-SAX Analysis Module which receives ECG signals acquired by the data acquisition module and converts, clusters and classifies them and (iii) Visualization Module which generates a visual representation of ECG data. The division of framework into subcomponents allows applications to be designed and developed irrespective of other components. These self-contained modules can be “plugged in” to the core engine.

RA-SAX analysis module is divided into two main components. The first one is the ECG classification module which converts the incoming ECG data into SAX strings, clusters them and classifies them. This component is attached with the second section of the system which is resource-adaptation scheme that aims to intelligently control the classification parameters following the resource's state of the system.

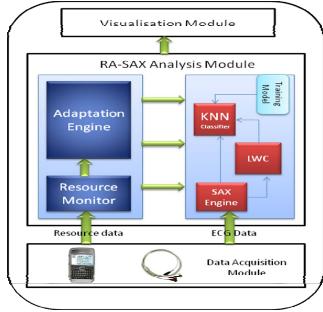


Figure 1: RA-SAX based mobile ECG analysis system architecture.

The key techniques in our framework are: *Resource-Aware* variation of *SAX* for representing the ECG signals as a time-series, *LWC (Light-weight Clustering)* for real-time data stream clustering on a mobile device, and *KNN* for classifying the ECG segments. All of these analysis components are managed by a resource-aware adaptation engine.

III. IMPLEMENTATION AND EVALUATION

We have developed a prototype of the Resource-Aware ECG Classification on the Android Developer phone using Eclipse as the development environment. The prototype was tested on Android Developer phone version 1.5. To obtain ECG signals, the Data acquisition module interfaces with SQLite to retrieve ECG values from a database that is stored on mobile device. SQLite is a light weight relational database supported on Android platform. For testing purposes, ECG values are stored as records on the device. The experiment setup consists of 29 abnormal samples and 18 normal samples. Every record has 512 values of ECG data containing recorded voltage in (mV).

In [4] we presented the experimental results pertaining to impact of resource-aware adaptation of varying the *segment size* and *alphabet size* parameters of the SAX representation. We recorded the change in classification time for each of the abnormal samples when number of segments were reduced from 300 to 100 while keeping the alphabet size constant. A conspicuous decline in the model building time is noticeable as the number of segments decreases. There is also a consequent but not very significant loss of accuracy. The second parameter identified for investigation was *alphabet size* which was changed to 20, 15 and 10. During the tests, changing the alphabet size showed negligible improvements on classification time. This was attributed to the fact that the classifier will still be using the same number of SAX characters to find the k nearest neighbors.

In this paper, we studied the effect of number of segments on battery drain time. In order to effectively perform resource adaptation, the battery drain time was noted for each of the test cases when the number of segments is set to 300, 200 and 100. We were able to show that when the number of segments is changed from 300 to 100, the device can function for an extra ($140 - 118 = 22$) minutes.

The third parameter we investigated was the clustering threshold. Varying the threshold of *LWC* is expected to result

in different accuracy for the clustering algorithm. Our experiments show that 2000 as the threshold (when varied between 1000 and 2400) gives the highest accuracy of 76% in our context.

One of the matters of investigation was comparing the framework with and without clustering algorithm. We conducted our experiments to measure time taken to classify in both cases while keeping segments to 300 and alphabet size to 26. We were able to show that the classification time is improved by up to fifty percent with the addition of clustering as shown in Figure 2.

The second set of experiments examined the accuracy of the classification module with and without clustering algorithm. The classification with clustering is promising in some cases and proves inaccurate in some other ones whereas the original module with no clusterer shows more consistency. In the best cases, the classification accuracy with clustering slightly improves while it can reduce up to 16% in worst cases.

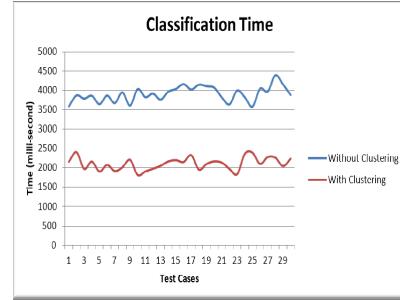


Figure 2: Classification time with and without clustering

IV. CONCLUSION AND FUTURE WORK

Our resource-aware architecture for mobile ECG analysis uses SAX, Light Weight Clustering (LWC) and KNN. The notion of having a clustering algorithm showed promise as a platform to conduct further studies. Improvements in time for classification suggest appropriate use of the module could save time, battery life and increase system reliability. In future, we intend to investigate the use of other clustering algorithms to identify the most suitable clustering algorithm for the context.

REFERENCES

- [1] J. J. Oresko, J. Zhanpeng, C. Jun, H. Shimeng, S. Yuwen, H. Duschl, and A. C. Cheng, "A Wearable Smartphone-Based Platform for Real-Time Cardiovascular Disease Detection Via Electrocardiogram Processing," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, pp. 734-740, 2010.
- [2] C. Bill, K. Eamonn, and L. Stefano, "Probabilistic discovery of time series motifs," in *9th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, Washington, DC, USA, 2003, pp. 493 - 498.
- [3] M. Gaber, S. Krishnaswamy, and A. Zaslavsky, "On-board Mining of Data Streams in Sensor Networks," in *Advanced Methods for Knowledge Discovery from Complex Data*, L. Jain and X. Wu, Eds.: Springer London, 2005, pp. 307-335.
- [4] Sinha, A., Tayebi, S., Krishnaswamy, S., Waluyo, A. B., and Gaber, M. M., (2011), "Resource-Aware ECG Analysis on Mobile Devices", Proc. of the ACM Symposium on Applied Computing – Data Streams Track, Tawain, Mar 22-25, ACM.