Computer Networks 136 (2018) 119–127

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Multi-level sample importance ranking based progressive transmission strategy for time series body sensor data

Ming Li, Joseph Reeves, Carlos Moreno*

Department of Computer Science, California State University, Fresno, Fresno, CA, USA

ARTICLE INFO

Article history:
Received 12 October 2017
Revised 27 January 2018
Accepted 5 March 2018
Available online 7 March 2018

Keywords:
Data analysis
Quality of services
Progressive transmission
Time series data
Body area networks

ABSTRACT

Body sensors have gained increasing interest during the past several years. With more applications deployed, it is imperative to ensure the success of data analysis, which largely depends on data transmission reliability as well as the importance of samples received. Traditional approaches focus on improving data reliability through various schemes such as the prioritization of MAC access. However, in this paper we analyze the characteristics of time series body sensor data and propose to rank sample importance based on a multi-level approach. With this approach, samples are grouped into five levels, indicating their importance with regard to data analysis. Then, a progressive transmission strategy is designed to transmit samples in order of their importance so that the overall received data quality is maximized. Extensive experiment results indicate that as much as 40–60% bandwidth saving can be achieved while meeting the requirements of data analysis algorithms.

1. Introduction

Body sensor networks (BSNs [1]) or wireless body area networks (WBANs [2]) have gained increasing interest during recent years. Key applications that may benefit from this new technique include remote health monitoring, sports training and entertainments. The ultimate objective of most such applications is to perform data analysis for classification, recognition and pattern prediction. Therefore, it is important that sufficient data quality is provided so that the accuracy of data analytics can be supported.

There are three primary factors affecting body sensor data quality: (i) limited bandwidth (around 240 kbps); (ii) network loss common for wireless networks; and (iii) size of the network. If there are only a couple of sensors, it is usually not a concern. However, with more deployment for patients with long care needs for multiple illnesses, how to guarantee quality of service (QoS) for multiple body sensors (such as ECG, EMG, EEG, etc.) remains a challenging issue. Existing approaches to address network loss include ensuring QoS support for specific data streams as well as increasing network throughput. These strategies, although effective in general, have not fully considered the unique characteristics of time series body sensor data.

Basically, the accuracy of data analysis, under the scenario of wirelessly transmitted sensor data, depends on two factors: (1) network delivery ratio, i.e., how many samples are successfully delivered through the network to the receiver; this is, in turn, determined by network capacity, channel reliability and traffic load; and (2) data recovery ratio, i.e., among all lost samples, how many of them can be recovered to a level that approximates the original sample. While network delivery ratio can be largely improved by QoS strategies, data recovery has to be addressed by closely looking at data characteristics and specific analytics algorithms.

An initial investigation of time series body sensor data may give the conclusion that data samples are largely independent and there do not exist structures similar to other types of data, e.g., I, B, and P frames in MPEG videos [3] or base and refinement layers in 3D meshes [4]. However, when it comes to the data recovery, we identified some samples that are considered more important, i.e., it is more difficult to obtain a close approximation of the original data if these samples are lost. Based on the level of difficulty, we can group samples to multiple levels based on the characteristics of points with curvatures formed by sample values, which we call “importance ranking”. Further evaluation of the ranking reveals that the first two levels, especially the first level, contribute over 90% of the data quality. Even if we only receive level 0 samples, the average deviation of recovered data is less than 10%. However, such data only consists of around 40–60% of the total data samples.

Based on importance ranking, we propose to implement a progressive transmission scheme that transmits samples of different importance levels through reliable or unreliable channels in order

* Corresponding author.
E-mail addresses: mingli@csufresno.edu (M. Li), josephreeves@mail.fresnostate.edu (J. Reeves), mmxzbnl@mail.fresnostate.edu (C. Moreno).

https://doi.org/10.1016/j.comnet.2018.03.001
1389-1286/

to meet the quality requirement of data analysis algorithms. Each segment of 200 samples is grouped into 5 importance levels, then packetized into priority data packets and non-priority packets. Priority packets are acknowledged by data aggregators with a maximum of 3 retransmissions to ensure its receipt. Non-priority packets are simply sent without ACK. Depending on the estimated accumulative data quality index (ADQI) by the aggregator and then a feedback for adaption, a sensor may decide to send more levels through priority packets and/or drop low importance level samples to save bandwidth.

A test bed has been implemented to study the performance of the proposed approach. The test bed consists of multiple TelosB motes running TinyOS [5] and a PC. One mote is attached to the PC so that it functions as the data aggregator through serial communication. Due to the limitation of TelosB sensors, we used both EMG data from PhysioNet [6] and accelerometer data collected with the BSN Development Kit v3 [7], which contains three types of Activity of Daily Living (ADL): jogging, sitting, and walking. Extensive experiments have been conducted with sensor data collected from real sensors. Results show that the first two levels consist of around 40–60% of the total data samples but they achieve a quality index of 90%, therefore achieving significant bandwidth savings while meeting the requirement of most data analysis. In addition, the progressive transmission scheme performs well under various loss scenarios.

2. Related works

QoS scheduling for streaming body sensor data has been well investigated. Zhou et al. [8] proposed BodyQoS, a virtual MAC for QoS scheduling in BSNs. The approach measures the effective bandwidth and adaptively allocates remaining resources to meet the QoS requirements of applications. A desirable feature of BodyQoS is that it does not require the modification of the underlying MAC layer implementation. Fulford-Jones et al. [9] proposed CodeBlue, an ad-hoc infrastructure for emergent medical care. In this project, several types of body sensors (e.g., pulse oximeter, ECG, and EKG) are individually connected to Zigbee-enabled radio transmitters, which communicate with access points directly. Due to the ad-hoc architecture and the capability of self-organizing, CodeBlue yields scalability for network expansion and flexibility to connect various wireless devices. Jiang et al. [10] proposed CareNet, an integrated wireless environment used for remote health care systems. CareNet offers features such as high reliability and performance, scalability, security and integration with web based portal systems. High reliability is achieved by using a two-tier architecture.

Patel and Wang [11] discussed various challenges in body area networks. They identified requirements for several specific types of applications in terms of data rate, number of nodes, bit error rate, time active and battery lifetime. Then, they discussed the design requirements of the radio frequency wireless systems, including QoS, antenna design, PHY protocol design, MAC protocol design, energy efficiency and interoperability with other devices. They also surveyed existing technologies which met the requirements and argued that the acceptance and adoption of body sensor networks can only occur with standardization, which will reduce costs, increase user comfort with the technology and also increase the usefulness of the devices by increasing interoperability.

There have been considerable standardization efforts during recent years. Among many potential technologies, Zigbee and Bluetooth are the most widely deployed. Zigbee is a very low power, collision avoidance protocol optimized for lower power sensors. It has developed a health care specific protocol and is compliant with all IEEE 11073 devices as well as most other IEEE 802.15.4 [12] wireless devices. Bluetooth supports high-bandwidth and many existing devices, and has a health care compliant version defined. However, Bluetooth has very high power requirements and uptime for the radios. Bluetooth Low Energy [13] is a new proposed system from Bluetooth which will have lower energy requirements and still be interoperable with Bluetooth Classic, but details are still forthcoming at this time. Being aware of the unique requirement of supporting a wide range of applications by body area networks, an IEEE 802.15.6 [2] task force has been working on finalizing a WBAN standard. The new standard aims at providing flexible and configurable energy efficient MAC operations.

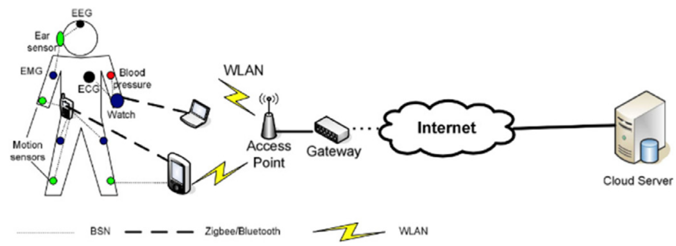

Fig. 1. System model/overview.

Progressive transmission tends to be dependent on the network bandwidth and available processing power and therefore has been used in multiple applications. Englert et al. [14] have proposed a streaming framework that attempts to be independent of the bandwidth and processing unit by utilizing a rating function to control priorities and a similarity metric to use already loaded units as replacements. Li et al. [4] proposed a progressive transmission scheme to reduce the performance gap between compression schemes and lossy network transmissions for 3D mesh streaming. This is done by utilizing a 3D middleware between the application layer and transport layer of triangle-based 3D models. The middleware is general such that it can handle multiple compression techniques, considers the end user's hardware capabilities and uses a dynamic cost selector to choose the transport protocol that works best for the current real-time network traffic. MobiTree [15] allows for 3D tree models to be transmitted to mobile phones, however at the cost of trading off latency and quality. CAPTURE [16] is a communications architecture for underwater relays that allows for image, sonar and time series sensor data to be progressively transmitted from an unmanned underwater vehicle to surface operators. Delport et al. [17] developed a wireless sensor network that is able to progressively transmit 2D image data. Finally, in [20], Englert et al. were able to perform progressive compression and transmission of 2D images over visual sensor networks.

3. System architecture, model and assumptions

In a typical medical application, multiple sensors such as ECG, EKG, EMG, EEG, motion sensors and blood pressure sensors send multimodal time series data to nearby data aggregators (cell phone, watch, headset, PDA, laptop, robot, etc.) based on the application needs. Then, through Bluetooth/WiFi, these data can be delivered remotely to the physician side for real-time diagnosis, or to a medical database for record keeping or to request for an emergency (Fig. 1). For such systems, the following assumptions are made:

- channel is unreliable and may drop data;
- with multiple body sensors communicating with the same aggregator, the channel may reach saturation status and therefore drop packets in the queue;
- reliable transmission of all samples incurs high protocol overhead for relatively small body sensor samples.

With strict requirements on the completeness of sensor samples by data analysis algorithms, missing samples at the data ag-