On the Effectiveness of Relaxation Theory for Controlling High Traffic Volumes in Body Sensor Networks

Naimah Yaakob[1], Ibrahim Khalil[1], and Jiankun Hu[2]

[1] School of Computer Science and Information Technology,
RMIT University, Melbourne VIC 3000, Australia
[2] School of Engineering and Information Technology,
University of New South Wales, Northcott Drive, Canberra
{yaakob.naimah,ibrahim.khalil}@rmit.edu.au, J.Hu@adfa.edu.au

Abstract. Congestion related issues are major concerns in any networking system including Body Sensor Networks (BSN). This is due to the number of disastrous effects (e.g. high packet loss rate and service interruption) it may cause on the system's performance. BSN, which normally involves with life-threatening measurements, are found to be very much affected by this problem. The incorporation of its real-time applications with life-death matters may likely put people at high risk during congestion. To address this challenge and alleviate congestion in BSN, we explore the feasibility of a new rate limiting technique known as Relaxation Theory (RT). Uniquely distinctive from the typical rate limiting schemes, the novelty of our approach lies in the ability to 'relax' or postpone the excessive incoming packets to a certain extent, and avoid congestion from occurring in the first place. An insight performance analysis on one of BSN applications in healthcare monitoring (Electrocardiogram - ECG) shows promising results.

Keywords: Congestion control, Relaxation Theory, Engineering Level.

1 Introduction

The limited buffer spaces in Body Sensor Networks have created substantial performance boundaries for the underlying applications [1]. With very limited capacities of bandwidth and storage, the buffer is highly likely to be overwhelmed by large volumes of sensors reading. This problem, which is known as *congestion*, has appeared to be even more challenging in BSN's domain applications that mainly involve with life-threatening measurements. Congestion may cause many other problems such as the increase in packet loss rate, high energy consumption and high service delay, all of which are detrimental to BSN's performance.

Fig. 1 shows BSN scenario in healthcare application and how congestion initiates. Heavy congestion may occur when multiple leads have data to be sent to the gateway or intermediate nodes, and sensors from several patients are sending their readings simultaneously. This will create high reporting rates that may

K.S. Nikita et al. (Eds.): MobiHealth 2011, LNICST 83, pp. 16–23, 2012.

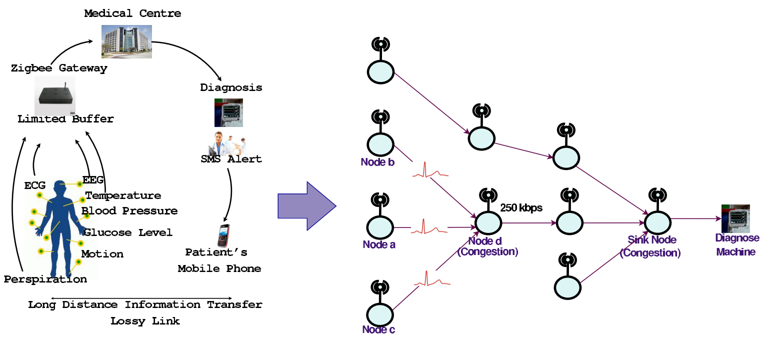


Fig. 1. BSN in Healthcare monitoring and congestion scenario

overload either the intermediate nodes or gateway's buffer and cause massive information loss. Ensuring reliable and timely packet delivery in healthcare applications are therefore of utmost importance to deliver accurate information. This might really help in avoiding false alarm that might lead to wrong diagnosis and affect the monitored patient. Late delivery and packet loss may result in obsolete data and thus inaccurate information. The loss or unreliable transmission of this kind of information might lead to patient's death, who would have survived should the packets arrived safely at the destination.

To overcome the above-mentioned problems, in this paper, we have introduced a novel scheme for managing and controlling high traffic volumes that can cause buffer overflow. This is done at an early stage of physiological data transmission at a sensor node using Relaxation Theory (RT) [2]. To the best of our knowledge, this is the first attempt to solve congestion issue directly associated with BSN. All the existing approaches [3–8] are not applicable for BSN due to its unique characteristics, colossal amount of traffics and different network topology. In addition, most of the existing techniques are mainly focusing on tackling the congestion after it has already occurred, while RT is designed to ensure that the system is steered clear from congestion in the first place. Our method avoids the formation of unfinished works in any nodes, hence ensures reliable packet delivery to the destination. Also, this method is able to eliminate the packet loss occurrence that have been major dilemma in BSN applications. The unfinished works, or sometimes termed as incomplete works, is defined as the number of excess packets still queued in the buffer at end of transmission. The term 'works' refers to packets and is used interchangeably in this paper.

By considering the healthcare monitoring system as shown in Fig. 1, our goal is to find an appropriate Engineering Level (EL), $L(f, \xi)$ [1] to ensure zero unfinished works and packet loss at the end of transmission. Note that the word

[1] The Engineering Level is defined as the number of packet slots in a combined pool of bandwidth, in order to be able to place the incoming packets, potentially after some buffering, without any additional loss or delay.

EL and $L(f, \xi)$ are used interchangeably to represent the Engineering Level. In particular, we aim at tackling the issue of losing important packets; either still queued in the buffer or lost due to congestion.

2 Related Research

The use of rate limiting in solving congestion issues has been widely explored in sensor networks as a fundamental approach to control high traffic volumes. While many of the proposed mechanisms [3–6] mainly focus on mitigating the congestion, we aim at preventing this problem at an early stage before its initiation.

The method found in [3] employed Additive Increase Multiplicative Decrease (AIMD) to adaptively control the sending rate based on the congestion notification from the child nodes. When the previous packet is successfully forwarded to destination, the intermediate node will increase the sending rate by a constant α. Otherwise, the sending rate is multiplied by the factor of β. In contrast, Event-to-Sink Reliable Transport (ESRT) [5] monitors the local buffer occupancies and notifies the source nodes to slow down the sending rates once buffer level exceeds certain threshold. Despite their success, these two methods did not considered the amount of incomplete works in the buffer at the end of transmission which is an important criteria for ensuring reliability in BSN. Another rate limiting technique known as Pump-Slowly, Fetch-Quickly (PSFQ) [6] also does not solve the packet loss problem due to congestion, hence is not feasible for BSN.

Our primary objective is basically similar to what is found in [7, 8]. However, in LACAS [7], the authors try to equate the packet arrival rate with packet service rate, the assumption that may not always hold for BSN. While this method can gracefully ensures congestion-free network in Wireless Sensor Networks (WSN), it might not be applicable to BSN. Huge excessive data over the service rate, might lead to high number of packets drop. On the other hand, CODA [8] also throttle nodes' transmission once congestion is detected. Though this technique can achieve energy efficiency, the reliability issue which is the main concern in healthcare application has not been addressed in this approach.

Different with those typical rate limiting approaches [3–8], our approach prevents congestion from occurring by deriving an appropriate $L(f, \xi)$ that constitutes together the service rate and buffer size of the system, so that the occurrence of packet loss due to congestion can be avoided in advance. Once the correct $L(f, \xi)$ is obtained, packets are either directly applied to the output pool, or kept in the specified buffer for at most ξ cell slots.

3 Relaxation Theory (RT): Concept and Overview

The main objective behind RT [2] is to find the EL value that can adhere and satisfy with the defined Quality of Service (QoS), so that the arriving packets can be 'relaxed' or 'postponed' to any value of ξ without additional loss or delay (still within the pre-defined period). Once the EL is properly defined, there will be a fairly allocated bandwidth to each packet without any contention. In reference to

Table 1. Mathematical Notations

No	Symbol	Explanation
1	ξ	Allowable unit delay
2	$L(f, \xi)$, h	Engineering Level
3	$f(t)$	Packets arrival function
4	$B = h\xi$	Buffer size
5	c_n	The point on x-axis where the average intersects with $f(t)$
6	$O(Xn)$	Accumulative unfinished works at specific time t
7	$O(b)$	Unfinished works at end of transmission time b

healthcare monitoring, avoiding contention among the sensed packets is crucial in ensuring packets loss from occurring and avoid misleading interpretation of the data. All the notations used can be listed in Table 1.

3.1 The Constraints

All the related constraints in obtaining the correct EL are defined as follows:

1. $O(b) = 0$
2. $Loss = 0$
3. $O(X_n) <= h\xi \;\; for\ all\ n$
4. $\xi < b - a$

The first and second constraints are important criteria for real-time healthcare monitoring system. In fact, BSN itself is very sensitive to packet loss [1] as this might trigger wrong diagnosis that may harm the critical ailing patients. The third constraint requires that the incomplete works at any time between a and b should be less than the buffer size ($h\xi$). This is to avoid the formation of buffer overloading. The largest unfinished works should occur at point c_n in such a way that $O(c_n) = h\xi$. The last constraint implies that the unit delay ξ should be less than $b - a$, or otherwise, the system may require a very large buffer to gracefully discharged all the packets, which is very impractical in scarce resources BSN.

3.2 Engineering Level (EL), $L(f, \xi)$

There are four parameters that need to be known in advanced in order to achieve the required $L(f, \xi)$ which are: Packets arrival function ($f(t)$), start time (a), end time (b), and allowable unit delay (ξ). Based on these parameters, the correct $L(f, \xi)$ can be derived as follows:

$$h = L(f, \xi) = max[h1, h2.., hn] \tag{1}$$

where h_n are various EL candidates. The h_n value is derived from the average arrival rate (A) given as $A(f, b-a) = \frac{1}{b-a}\int_a^b f(t)\,dt \;\; = 0$. Then the $L(f, \xi)$ can be obtained as follows:

Table 2. Experimental Setup

No	Input Parameters	Setup
1	Bandwidth	250 kbps
2	Arrival Rates	1-12 leads
3	Packet Size	30 bytes
4	ξ	1, 2, 3
5	Sensor Leads	1-12

$$L(f,\xi) = max[max[\frac{1}{b-a}\int_a^b f(t)\,dt, \frac{1}{b-c_n}\int_{c_n}^b f(t)\,dt], \frac{1}{c_n-a+\xi}\int_a^{c_n} f(t)\,dt] \quad (2)$$

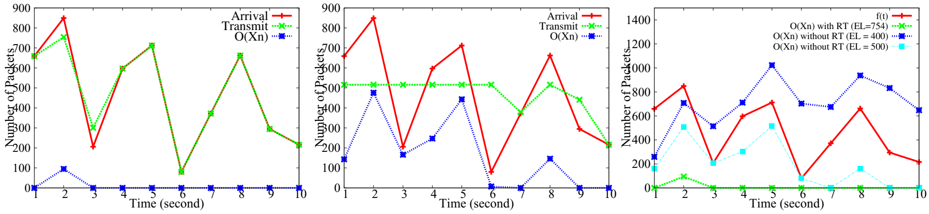
Value c_n is obtained from the intersection point between (A) and the x-axis. Note that $O(c_n)$ is the largest unfinished works. Finally, we need to check whether $O(c_n) = h\xi$ and $O(b) = 0$, which are the two significant validation tests to know that the obtained $L(f,\xi)$ is absolutely correct. In other words, the incomplete works that exceed the buffer size at any time in the interval, may violate the rules and thus may result in some packet loss.

3.3 Experimental Setup

We tested the potentiality of RT using some numerical analysis which have been conducted based on some configuration setup as specified in Table 2. It is suffice to mention that this experiment is conducted based on the assumption that the data are collected from only one patient at this point. We believed that controlling the incoming traffics and managing the buffer at this early stage are the key success of the congestion control in the entire system. If the buffer can be emptied at each sensor node including the intermediaries, the congestion can highly likely be alleviated from the whole network system. For the purpose of this experiment, we generate random number of packets using the Random Number Generator (RNG). This has been carried out in C language in order to create different arrival rates at different time unit. Once the incoming packets have entered the input trunk, RT is immediately applied and the system will check for the available slots of transmission based on the calculated $(L(f,\xi))$.

4 Results and Analysis

This section demonstrates some parts of the results obtained through the numerical analysis. The performance have been evaluated using two performance metrics: Number of unfinished works at time b, $(O(b))$ and Number of Packet Loss.



(a) ξ = 1, EL = 754 (b) ξ = 2, EL = 516 (c) ξ = 1

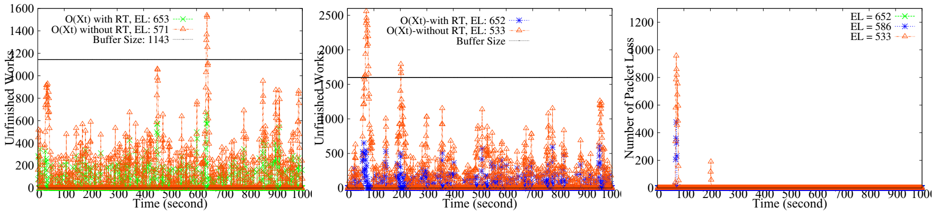
Fig. 2. The resulting $O(Xn)$ with the calculated EL for different ξ values (a) $EL = 754$ (b) $EL = 516$ (c) Performance comparison with and without RT

Fig. 2a shows the resulting number of $O(Xn)$ for $\xi = 1$. From the figure, we noticed that there is no unfinished works at all time except once while $t = 2$. The calculated EL for this case is 754. Thus, if $f(t)$ at any time t are less than the EL, they will be transmitted as per arrival. At $t = 2$, there is excessive arrival of 95 packets beyond the available EL slots. These extra packets are therefore stored in the buffer. Since the buffer is still empty and not yet being occupied, at $t = 2$, there are enough spaces to keep the packets until the next transmission. In this case, the buffered packets can be delayed to the future until one unit of time.

At $t = 3$, there are only 206 packets arrived, thus there are ample of empty spaces to accommodate together those stored in the buffer. That is explained by the transmission line that is slightly above the packets arrival in Fig. 2a at that particular time. As such, there is no unfinished works left ($O(b) = 0$) which validates the finding of EL. On top of that, the absence of packets loss at any time t, further verifies the obtained results.

We extended the value of ξ to 2 and 3 and the results obtained are as depicted in Fig. 2b. To our surprise, the results for $\xi = 2$ and 3 are absolutely the same. This is because, the shape of the graph is determined by the obtained EL, which is the same for both the cases. We have a strong belief that this scenario is created when the obtained value of $h2$ is always higher than $h1$. Hence, no matter what ξ is, we will always get the same EL value.

The EL for this case is 516. Since EL value is far below almost half of the arrival function $f(t)$, there are buffered packets as early as $t = 1$ until $t = 6$. All the packets stored in the buffer, are then being transmitted together with the current arrival of packets, and sent at the maximum of $(EL - f(t))$. At $t = 6$, there are still 443 $O(Xn)$ in the buffer. However, the sum of those in the buffer and the new arrival exceed the EL for about 7 slots. Hence, these extra packets will be kept for transmission in the next round. The processes are repeated until end of time b, whereby all of the works have been fully transmitted. Since the EL obtained in both cases satisfy all the defined constraints, no packets are lost or dropped as can be seen in the graphs. Therefore, the ELs have been properly selected.



(a) (b) (c)

Fig. 3. (a) Performance comparison between proper and improper selection of EL. $O(Xn)$ above the buffer size is considered lost. (b) Another performance comparisons for different sets of data (c) The resulting packet loss for (b) in different EL.

Performance comparison of $O(Xn)$ with and without RT is shown in Fig. 2c. Our method outperforms the one without RT all the times, and the resulting $O(b)$ is also 0 as evident in Fig. 2a and Fig. 2b. Therefore, the implementation of RT in this experiment has substantially improved the performance and eliminated the occurrence of congestion and packet loss in advance.

The significant contributions of zero unfinished works and packet loss presented in this paper have eliminated the issue of incomplete transmission of packets at the end of transmission, while gracefully solving the information loss problems. These contributions are crucially needed in providing and ensuring a reliable transmission in healthcare monitoring. The losses of vital signals in such real-time environment may affect the ongoing diagnosis and lead to false interpretation of the end result that may harm the monitored patients. Since the diagnosis may involve with life-death matters, careful handling of the collected readings is of the major concern.

As we extended the experiment further to $b = 1000$, the number of $O(Xn)$ happened to appear more frequently since it is an accumulative incomplete works. However, with the correct derivation of EL, these $O(Xn)$ tend to be vanished as the time approaches the final point b. This is depicted in Fig. 3a.

In comparison with that, the value lower than the calculated EL (in this case is 571) produced high number of $O(Xn)$ to the extent that violate the RT rules at some points (points that exceed the buffer size) and results in the number of packet loss due to improper selection of EL. Another RT performance comparison is shown in Fig. 3b. This new data set possess a different EL value (652). Again, two different values (586 and 533) of lower rates have been chosen for comparison purposes. As expected, the correct EL gives outstanding performance compared to the other two. As discussed earlier, BSN performance decreases without the use of RT. This can be obviously noticed in the graph. In fact, some of the packets have surpassed the limited capacity of the buffer, and thus lost as depicted in Fig. 3c. These information loss may be detrimental to BSN operation as it might contain useful knowledge. On the other hand, no packet loss has occurred when the correct EL value was used.

All the results obtained in this experiment show promising features in assisting good BSN services. The outcomes have brought significant improvement to BSN since it is very sensitive to any information loss. The significant contribution of zero unfinished works presented in this paper has eliminated the issue incomplete transmission of packets (in the buffer) at the end of transmission, while gracefully solving the packet loss problems.

5 Conclusion and Future Works

In this paper, we have introduced a new rate limiting technique using an approach known as Relaxation Theory. Based on the preliminary analysis, we are able to demonstrate that the properly selected EL is capable of improving BSN performance and avoid congestion in advance. We have also highlighted several constraints that need to be followed in order to obtain the desired EL. The preliminary results demonstrate promising features for our future experimentations in solving many congestion-related issues in BSN. As part of our future works, we would also be interested to study the RT performance in larger network size and investigate the performance trade-off in terms of delay and buffer size.

References

1. Yang, G.Z.: Body sensor networks, research challenges and applications
2. Minoli, D.: Broadband Network Analysis and Design. Artech House, Inc., Norwood (1993)
3. Woo, A., Culler, D.E.: A transmission control scheme for media access in sensor networks. In: Proceedings of the 7th Annual International Conference on Mobile Computing and Networking, MobiCom 2001, pp. 221–235. ACM, New York (2001), http://doi.acm.org/10.1145/381677.381699
4. Stann, F., Heidemann, J.: Rmst: reliable data transport in sensor networks. In: Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications, pp. 102–112 (May 2003)
5. Akan, O., Akyildiz, I.: Event-to-sink reliable transport in wireless sensor networks. IEEE/ACM Transactions on Networking 13(5), 1003–1016 (2005)
6. Wan, C.-Y., Campbell, A., Krishnamurthy, L.: Pump-slowly, fetch-quickly (psfq): a reliable transport protocol for sensor networks. IEEE Journal on Selected Areas in Communications 23(4), 862–872 (2005)
7. Misra, S., Tiwari, V., Obaidat, M.: Lacas: learning automata-based congestion avoidance scheme for healthcare wireless sensor networks. IEEE Journal on Selected Areas in Communications 27(4), 466–479 (2009)
8. Wan, C.-Y., Eisenman, S.B., Campbell, A.T.: Coda: congestion detection and avoidance in sensor networks. In: Proceedings of the 1st International Conference on Embedded Networked Sensor Systems, SenSys 2003, pp. 266–279. ACM, New York (2003), http://doi.acm.org/10.1145/958491.958523