Relocation of Gateway for Enhanced Timeliness in Wireless Sensor Networks

Kemal Akkaya and Mohamed Younis
Department of Computer Science and Electrical Engineering
University of Maryland, Baltimore County
Baltimore, MD 21250
kemal1 | younis@cs.umbc.edu

Abstract

In recent years, due to increasing interest in applications of wireless sensor networks that demand certain quality of service (QoS) guarantees, new routing protocols have been proposed for providing energy-efficient real-time relaying of data. However, none of these protocols considered any possible movement of the sink node for performance purposes. In this paper, we propose possible relocation of sink (gateway) for improving the timeliness of real-time packets. Our approach searches for a location close to the most loaded node. The gateway is then relocated to the new location so that the load of that node is alleviated and the real-time traffic can be split. As long as the gateway stays within the transmission range of all last hop nodes, it can be moved to that location without affecting the current route setup. Otherwise routes are adjusted by introducing new forwarders. Simulation results demonstrate the effectiveness of the proposed approach.

1. Introduction

Networking unattended wireless sensors are expected to have significant impact on the efficiency of many military and civil applications such as combat field surveillance, security and disaster management [1]. The unique characteristics of sensor networks have made efficient routing of sensor data one of the technical challenges in wireless sensor networks. To address such challenges, the bulk of research on routing in wireless sensor networks mostly aim at maximizing the lifetime of the network, allowing scalability for large number of sensor nodes and supporting tolerance for sensor's damage and battery exhaustion [2]. These performance objectives have been deemed sufficient for applications, which do not require on-time response or for which data are not collected at high rate.

However, there has been an increasing interest in sensor networks applications that require certain performance guarantees such as end-to-end delay. For instance, routing of imaging data in a battle environment requires careful handling in order to ensure that the end-to-end delay is within acceptable range and the images are received properly without any distortion. Other typical applications include real-time target tracking, emergent event triggering in monitoring applications and critical information relaying in emergency applications. Since most of the current protocols do not provide performance guarantees for such applications, new routing protocols that can achieve desired Quality of Service (QoS) for the delivery of sensing data are proposed. The proposed routing protocols not only ensure soft real-time delay guarantees through the duration of a connection but also provide the use of most energy efficient path [3][4]. While such protocols achieve soft end-to-end delay guarantees, their service can diminish with the increasing volume of real-time data. In such cases, most of the packets can start to miss their specified deadlines. In order to enhance timeliness in such situations, one of the solutions is to explore gateway's ability to move to a location where volume of real-time data is high. An example of this scenario is when the gateway is a laptop computer or other portable devices on the backpack of a rescue crew who is not expected to travel long distances. Such relocation in those circumstances can balance the traffic load among multiple nodes and hence increase the hit ratio of real-time packets.

In this paper, we present a novel mechanism for relocation of the gateway under QoS traffic for enhanced on-time delivery in sensor networks. End-to-end delay bound for real-time data is achieved through the use of a Weighted Fair Queuing (WFQ) based packet scheduling technique in each sensor node [5]. WFQ considers a different queue for each incoming flow and has been shown to provide, in statistical term, an upper bound on path delay for a leaky bucket constrained flow. When average hit ratio for real-time packets starts to decrease, our approach considers moving the gateway to a better position in order to maintain the same or even better level of timeliness. To determine the new gateway's position, we consider locations on or close to heavily loaded last hop node and try to split the incoming traffic passing through that node without extending the delay experienced by real-time packets over other routes. As long as the gateway remains within the transmission range of all the last hop nodes, we maintain the same routes that were set initially. If it is expected that the new location will put the gateway out of the transmission range of some of the last hop nodes in the current routes, new forwarder nodes that are not involved in any routing activity are selected. Such unused nodes will introduce very little queuing delay, which is desirable for timeliness of all real-time packets using those nodes as relays. If such nodes cannot be found, relocation is either not considered or a new network topology is set at the new location depending on the overhead rerouting will introduce relative to the gain in timeliness.

In the balance of this section we describe the sensor network architecture that we consider and summarize the related work. In section 2, we first summarize the underlying WFQ based real-time routing protocol and then describe our approach for handling issues related to the relocation of the gateway. Section 3 discusses simulation results. Finally we conclude the paper with a summary in section 4.

1.1. System Model

A set of sensors is spread throughout an area of interest to detect and possibly track events/targets in this area. A gateway node, which is significantly less energy-constrained than the sensors, is deployed in the physical proximity of sensors. The gateway is assumed to know the geographical location of deployed sensors. The gateway is responsible for organizing the activities at sensor nodes contingent to achieving a mission, fusing data collected by sensor nodes, coordinating communication among nodes and interacting with command node. While sensor nodes are stationary, we are considering a limited mobility model for the gateway. Sensors are assumed to be within the communication range of the gateway node. The sensor is assumed to be capable of operating in an active mode or a low-power stand-by mode. The sensing, processing and radio circuits can be powered on and off. In addition, the radio's transmission power is assumed to be programmable for a required range. It is worth noting that most of these capabilities are available on some of the advanced sensors, e.g. the Acoustic Ballistic Module from SenTech Inc. [6].

1.2 Related Work

While contemporary best-effort routing approaches address-unconstrained traffic, very little research has been done on QoS routing in wireless sensor networks. The first protocol for wireless sensor networks that includes the notion of QoS in its routing decisions is the Sequential Assignment Routing (SAR) [1]. Another QoS routing protocol for sensor networks that provides soft real-time end-to-end guarantees is SPEED [3]. SPEED strives to ensure a certain speed for each packet in the network so that each application can estimate the end-to-end packet delay by dividing the distance to the sink by the speed of the packet before making the admission decision.

The approach of [7] finds a least cost and energy efficient path that meets end-to-end delay requirements. The link cost used is a function that captures the nodes' energy reserve, transmission energy, error rate. In order to support both best effort and real-time traffic at the same time, a class-based queuing model is employed. The queuing model allows service sharing for real-time and non-real-time traffic.

All of the described protocols assumed a stationary gateway and did not consider relocation of the gateway node for enhanced performance. The only work to consider relocation of gateway node in order to reduce average energy per packet and hence increase the lifetime of the network in wireless sensor networks is reported in [8]. While the presented approach aims at optimizing the energy consumption and the lifetime of the network, it works only for unconstrained data traffic and does not fit to the needs of real-time traffic. Our approach is an extension to that work which works under the involvement of real-time traffic as well. It periodically checks for possible performance degradation and triggers relocation for the gateway node when needed. To the best of our knowledge, our work is the first to introduce relocation of gateway for enhanced end-to-end delay and miss ratio when real-time traffic is involved.

2. Relocation of Gateway

In multi-hop wireless networks, the throughput, average delay and energy consumed in packet routing depends on the positions of the sources of the data and the destination. The location of the gateway becomes very influential when real-time traffic is involved. For instance, in some circumstances most of the real-time path establishment requests can be denied or the hit ratio for real-time packets can decrease significantly due to congested routes through the gateway. Traffic congestion can be caused by the increasing the number of real-time data packets coming from nodes close to a recently event. In such circumstances, it may not be feasible to meet the requirements for real-time data delivery. This scenario can be unacceptable for applications like disaster management and target tracking, since failing to provide real-time traffic within the required time period may negatively impact the application. If the gateway has a capability to move, it can be beneficial to relocate the gateway close to the occurring event in order to spread the traffic on more hops and increase the feasibility of meeting the real-time delivery requirements. In this section, we first describe the underlying routing mechanism for providing on-time delivery of real-time data and then explain our approach of when, where and how to relocate the gateway for enhanced performance.

2.1 Real-time Routing Mechanism

Our routing approach for supporting real-time traffic is based on the model defined in [9]. Each node employs a packet scheduling discipline that approximates Generalized Processor Sharing (GPS). Since WFQ is flow based, the approach of [4] uses an approximation of WFQ by considering each imaging sensor node as a source of different real-time flow with only one real-time queue to accommodate the real-time data coming from these multiple flows (Fig. 1). The service ratio "*r*" is the bandwidth ratio set by the gateway and is used in allocating the amount of bandwidth to be dedicated to the real-time and non-real-time traffic on a particular outgoing link. This value is also used to calculate the service rate for each type of traffic on that particular node, with $r_m \mu$ and $(1- r_m) \mu$ being respectively the service rate for real-time and non-real-time data on sensor node m. In this case, r_m for the real-time queue on a node is the summation of link shares of all real-time flows passing through that node as shown in Fig. 1. In order to setup real-time routes to the gateway, the approach first obtains a set of energy-efficient paths for each real-time source. Then, these paths are further checked to identify the one that can meet the end-to-end delay requirement by trying to find an *r*-value for each node on that path. Once an r_m between 0 and 1 can be found for each node, the real-time connection is established for that path.

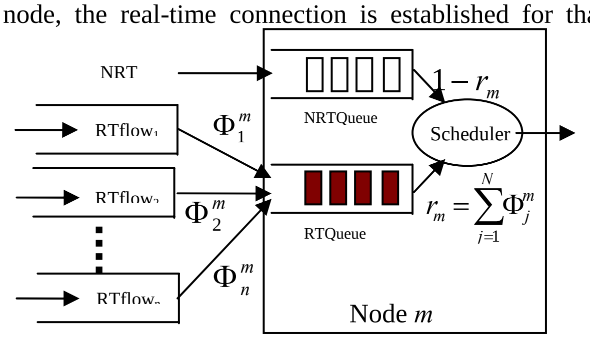

Fig. 1: Queuing model on a node

2.2 Relocation Approach

There are three issues to consider in gateway relocation. These are when to relocate the gateway, where to put it and how to handle its motion without disrupting data traffic. In this subsection, we will discuss these issues and then state our algorithm to handle such issues.

When to move the gateway: First of all, the time for such relocation should be decided based on some performance observations such as miss ratio. The miss ratio is indirectly related to finding proper *r*-values in our underlying approach. Note that, our approach for real-time routing basically tries to find an *r*-value by considering every path from the source of real-time traffic to the gateway. In cases where a proper *r*-value between 0 and 1 cannot be found, the connection is simply rejected and the path is not established. Moreover, even proper *r*-values are found, the miss ratio can start to increase due to involvement of new events that cause generation of additional real-time packets and hence more congestion. Therefore, the trigger for relocation will be the unacceptable increase in the miss ratio of real-time packets. In our approach, the gateway will set a threshold for the maximum level of miss ratio by maintaining such statistics periodically and consider relocating the gateway to a better location when such threshold is reached. Note that, the gateway can be relocated more than once whenever necessary during the data traffic cycles.

Where to move the gateway: After deciding that the gateway is to be relocated, our approach will consider searching a new location for the gateway. The main aim here is to move the gateway towards the loaded nodes in terms of real-time traffic so that the end-to-end delay can be decreased. Therefore, our approach first searches the last hop nodes i.e. the nodes that are directly transmit to the gateway in order to designate the hop with the biggest *r*-value and consider relocating the gateway at the position of that hop. This will helpful in twofold: (1) It will help in decreasing the average end-to-end delay since the number of hops for data packets to travel will be decreased. (2) It will help in admitting more real-time flows since the load is alleviated by splitting the traffic.

However, finding such new location for the gateway does not necessarily mean that the gateway will be relocated. Before moving the gateway, an analysis should be performed in order to assess the potential overhead that will be introduced when gateway is relocated at the new location. Such overhead will be due to any necessary route adjustment when the gateway goes out of the transmission range of some of the other last hop nodes. If this is the case, new relay nodes should be found to maintain uninterrupted data delivery and consequently some routes will need to be updated accordingly. While involvement of such new relay nodes will introduce additional delay for real-time packets and increase energy usage, forming new network topology based on the new location of the gateway can be undesirable due to its overhead in terms of control traffic and energy.

Therefore, when a new location is found, our approach first checks whether the gateway will go out of range or not at the new location. This is also used for breaking the ties when multiple alternative nodes with the same *r*-value are found. In such circumstances, if the other last hop nodes can still reach the gateway by increasing their transmission power, relocation can be performed safely. A pictorial illustration of this situation is depicted in Fig. 2a and 2b. If we assume that there is an occurring event on north-west of node A, many of the imaging sensors that are sources of real-time traffic will be turned on in that region, increasing the traffic density flowing through node A (Fig. 2a). In this case, most of the paths passing through node A will be rejected since the limited service rate on node A increases the end-to-end queuing delay. However, the load on node A can be alleviated by moving the gateway to the location of A. If the gateway is still reachable by the nodes B and C, they just increase their transmission ranges for uninterrupted data delivery as seen in Fig 2b.

If the new location requires some topology changes then our approach considers alternative positions that might be feasible for the gateway to be positioned at without causing any effect on the network topology. Such alternatives include the location of one of the siblings of the node with the biggest *r*-value or some other location that will be closer to that sibling and enable data relaying without affecting the latency of other real-time packets. Here the selection of the sibling node is decided based on its *r*-value since targeting the sibling with the biggest *r*-value can help best in splitting the real-time traffic. An example of this situation is shown in Fig. 2c. In this case, A_1, which is the sibling of A with the biggest *r*-value, is picked and gateway is relocated closer to that node so that the load of A is split between A and A_1. Such relocation does not require any route updates for the current topology. Note that, here the new position of the gateway should be within the communication range of A_1 and the other nodes A, B and C so that the traffic passing through A_1 can be directed to the gateway to alleviate the load of A. At

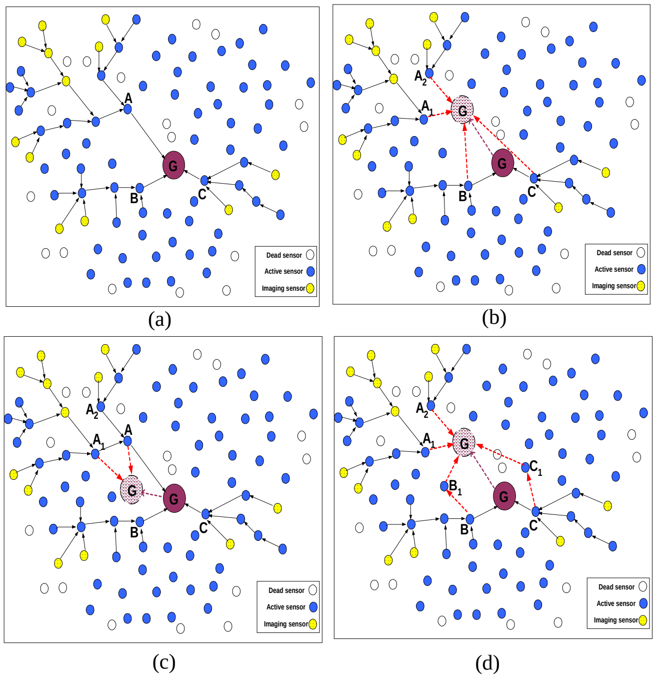

Fig. 2: (a) Initial routes (b) Gateway is relocated to the location of A if it is not out of range of C and B. (c) If that is not possible, gateway is relocated to a location close to B. (d) Otherwise unused forwarders B_1 and C_1 are found.

the same time nodes A, B and C can still reach the gateway without any need to adjust the routes.

In cases where a new location that will not cause the gateway to be out of range cannot be found, our approach strives to minimize the overhead of adjusting the routes. Therefore, new forwarder nodes that are currently not involved in real-time traffic relaying are picked for each of the last hop node in order to provide uninterrupted delivery of real-time packets. Since such nodes will not introduce extra queuing delay for real-time data, this will not affect the end-to-end delay of real-time data using that path. An example is shown in Fig. 2d. In this case, nodes B_1 and C_1 that are not part of any existing routes are designated as forwarders for nodes B and C respectively.

Failure to find such forwarders will result in keeping the gateway in its current position since the possible gain in relocation can be degraded due to overhead explained above. It should be noted that in some network architectures the gain achieved by the gateway relocation is valued more than the rerouting overhead. In such architectures it would make sense not to give up on gateway relocation if the network topology cannot be maintained. Since this scenario depends highly on the routing protocol and network management strategy, we have decided to flag such case for further architecture specific analysis or to simply seize the gateway movement.

How to move the gateway: The data transmission during the movement of the gateway is also a concern. It will be unacceptable to increase packet losses on some data paths. Our gateway motion handling mechanism tries to maintain continual packet delivery to the gateway through the adjustment of the transmission power of some sensors or via designating some forwarders to extend the current routes.

Once the new location is determined, the gateway explores two options based on the information of whether it will be out of range at the new location or not: If the gateway can still be reachable by the last hop nodes when relocated, it will simply instruct these nodes to adjust the transmission power of their radio to cover the gateway's new location and starts moving there. If the gateway detects that it would go out of the transmission range of last hop nodes and cannot receive the data from other relay nodes at the new location, it explores the option of employing sensor nodes to forward the packets. Ideal forwarder nodes should not be currently involved in relaying real-time traffic as stated earlier. The gateway starts moving to that location and assumed to move in strides to reach intermediate positions. The strides are to be in straight line. At each intermediate position, the gateway checks whether it can still be reachable by the last hops while traveling on the next stride and inform the last hop nodes about its situation. Once, it detects that forwarder nodes are needed, the routes are extended by those nodes and that information is sent to relevant nodes.

The proposed algorithm for gateway relocation is shown in Fig. 3. The gateway monitors the miss ratio for real-time data periodically to detect situations where there is need for relocation (line 1-2). If the gateway motion is justified, a new position is identified (line 3) and the gateway starts to move to its new location (line 4). In order to designate a new location for the gateway, the algorithm finds the node with the biggest *r*-value in line 6. In case of ties, the one closest to the gateway is picked in lines 7-8. Once the new location is found, the algorithm checks whether moving to that location will require employing forwarder nodes (line 9). If that is the case, alternative locations are searched (line 10). If no location is available without involvement of new forwarders, the algorithm looks for nodes that are not currently part of any data paths in line 12 and starts moving the gateway to the new position in line 15. When such forwarders cannot be identified, relocation is not performed (line 13-14).

3. Experimental Validation

```
Check the miss ratio at a predefined period
if (Miss ratio < Threshold)
        SearchForNewLocation()
Start moving gateway to NewLocation(G)
/* LastHopNodes: L */
SearchForNewLocation()
 Find the node s.t. m ← max_(j∈L) (rval^j)
 if (more than one m is found)
        Get the one closest to G
 if (G out of range of any of the nodes in L)
        Consider moving G closer to one of its siblings
        having the biggest r-val
        if (G still goes out of range of any node j∈L)
                FindForwarder(j)
                if (no forwarder can be found)
                    break; /* Do not relocate */
                NewLocation(G) ← Pos( m)
else NewLocation(G) ← Pos(m)
FindForwarder(j)
  Find Forwarder i  s.t.{Dist(j,i)<TRange[j] &&
  Dist(i,G_new)<TRange[i] && rval^i==0 &&
  Remain_energy(i) > δ }
 Update RouteTable[j];
 Update the list of L accordingly
```

Fig. 3: Relocation algorithm for gateway under QoS traffic

We have adapted the network operational model of [9] for validating our approach. In the experiments, the network consists of varying number of sensor nodes (50 to 200) randomly placed in a 500×500 meter square area. The gateway initial position is determined randomly within the region boundaries. A free space propagation channel model is assumed [10] with the capacity set to 2Mbps. Each node is assumed to have an initial energy of 5 joules. A node is considered non-functional if its energy level reaches zero. The maximum transmission range for a sensor node is assumed to be 50 meters [11]. For a node in the sensing state, packets are generated at a constant rate of 1 packet/sec. This value is consistent with the specifications of the Acoustic Ballistic Module from SenTech Inc. [6]. The sources generating data are assumed to be leaky bucket constrained with the maximum burst parameter σ of 10 packets. We assume that the network is tasked with a multi-target tracking mission in the experiment and the gateway can move with a maximum speed of 5m/sec when needed. We also assume that each sensor node is capable of taking the image of target to identify it clearly and can turn on its imaging capability on demand. The imaging sensor set may change with the movement of the target. The packet generation rate for imaging sensors is bigger than the normal sensors. Packets, generated by imaging sensors, are labeled as of real-time type and treated differently at the relaying nodes. The *r*-value is initially assumed to be 0 but it is recalculated as imaging sensors get activated. The default end-to-end delay requirement for real-time data is taken to be 0.08 sec [12]. Given our interest in enhancing on-time delivery of real-time data by employing gateway relocation, we used *average delay per packet, deadline miss ratio, average energy per packet, network throughput* metrics to capture the performance of our approach.

3.1 Performance Results

In this section we present some performance results obtained through simulation. As a baseline approach, we have used the same underlying routing mechanism with a stationary gateway for i.e., without considering any movement.

Delay and Timeliness: When qualifying the impact on the end-to-end delay and miss ratio for real-time packets we have observed that the relocation approach significantly decreases average delay per real-time packet and provides at least 20% decrease in the deadline miss ratio as shown in Fig. 4 and 5. This is due to the decreased number of hops and queuing delay for the data coming from highly loaded areas when gateway goes closer to those areas. The decrease in the miss ratio is even more significant when the number of sensors is increased, suggesting the positive effect on scalability.

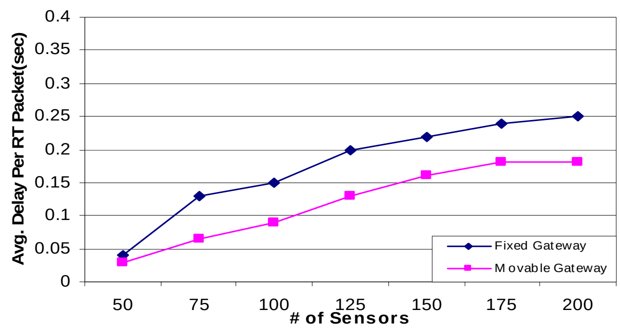

Fig. 4: Average delay per real-time packets with different number of sensors.

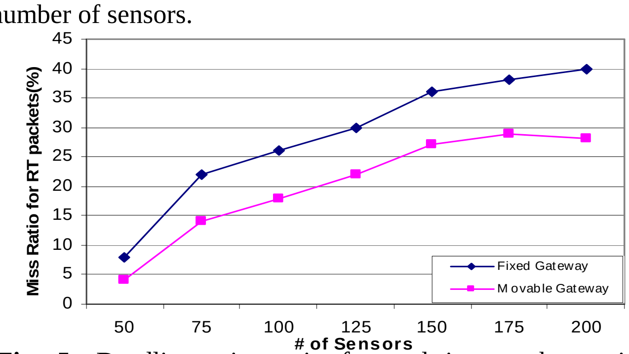

Fig. 5: Deadline miss ratio for real-time packets with different number of sensors.

Energy Consumption: In order to capture the effect of relocation on energy consumption, we have looked at the average energy per packet. The results, depicted in Fig. 6, demonstrate that our approach does bring very negligible energy overhead to the network. Although either some of the last hop nodes increase their range or routes are extended through additional forwarders, the potential energy overhead due to such adjustments is compensated through the energy gain when gateway is relocated towards loaded nodes.

Throughput: When we have looked at the throughput for real-time data, we have observed that relocating the gateway increases real-time data throughput by about 20% compared to the fixed gateway real-time data throughput, as shown in Fig. 7. This is expected since moving the gateway towards

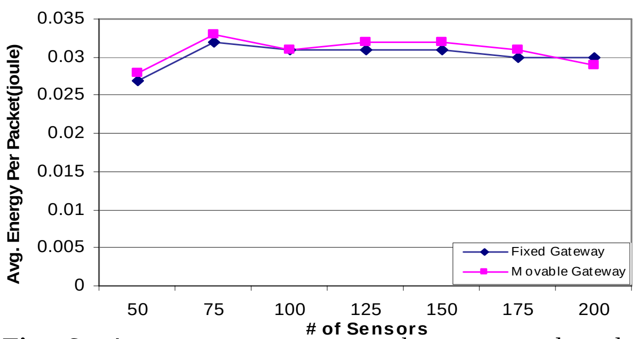

Fig. 6: Average energy per packet consumed under different number of sensors.

heavily loaded nodes and splitting the traffic will help in boosting the number of admitted flows for real-time relaying.

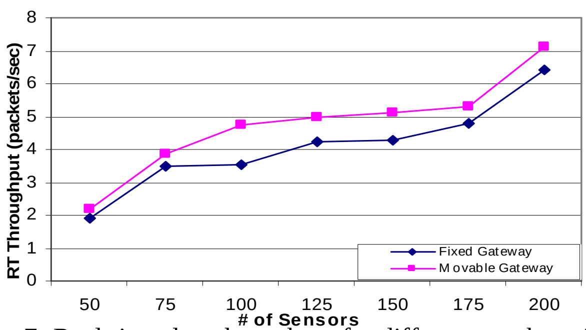

Fig. 7: Real-time data throughput for different number of sensors.

In summary, the simulation results clearly confirm the effectiveness of our approach. The gateway relocation under real-time traffic can provide substantial enhancement in hit ratio of real-time packets without introducing any extra overhead in terms of energy. Such enhancement is more significant at high traffic loads. Moreover, the throughput for real-time data has significantly benefited from such a move.

4. Conclusion

In this paper, we have presented an algorithm for effective gateway relocation in order to enhance the performance of routing QoS traffic in wireless sensor networks. At each sensor node, WFQ-based packet scheduler is employed. WFQ allows service sharing for real-time and non-real-time flows and achieves certain delay bounds when used along with leaky bucket constrained sources. Our approach periodically checks the miss ratio for real-time packets and triggers a relocation stimulus for the gateway if the miss ratio is above a certain threshold. In order to designate a new gateway location, our approach finds the node that routes the largest number of real-time packets and checks whether moving to that location or close to that location affects the current routes or not. Relocation is performed only if the new location will decrease the end-to-end delay for most real-time packets without increasing the energy usage of the network.

Simulation results have demonstrated the effectiveness of our approach showing at least 20% decrease in the deadline miss ratio and real-time data throughput when compared to a fixed gateway model. Moreover, the simulation has indicated that our approach boosts these performance metrics without major negative impact on the consumed energy.

References

[1] I. F. Akyildiz et al., “Wireless sensor networks: a survey”, *Computer Networks*, Vol. 38, pp. 393-422, 2002.

[2] K. Akkaya and M. Younis, “A survey on routing protocols for wireless sensor networks,” in *Elsevier Journal of Ad Hoc Networks* (to appear).

[3] T. He et al., “SPEED: A stateless protocol for real-time communication in sensor networks,” in the *Proceedings of IC DCS*, Providence, RI, May 2003.

[4] K. Akkaya and M. Younis, “Energy-aware delay constrained routing in wireless sensor networks,” Technical Report, Dept. of Computer Science, UMBC, 2003.

[5] A. K. Parekh and G. Gallager, “A generalized processor sharing approach to flow control in integrated services networks Services Networks: The Multiple Node Case”, in the *Proceedings of IEEE INFOCOM’93*, Vol. 2, pp. 521-530.

[6] "Data sheet for the Acoustic Ballistic Module", SenTech Inc., http://www.sentech acoustic.com/

[7] K. Akkaya and M. Younis, “An Energy-Aware QoS Routing Protocol for Wireless Sensor Networks,” *Proc. of the IEEE MWN 2003*, Providence, RI, May 2003.

[8] M. Younis et al., “Base-Station Repositioning For Optimized Performance of Sensor Networks,” in the *Proceedings of the IEEE VTC 2003-*, Orlando, FL, October 2003.

[9] M. Younis, M. Youssef, K. Arisha, “Energy-Aware Routing in Cluster-Based Sensor Networks”, in the Proceedings of IEEE/ACM MASCOTS2002, Fort Worth, TX, October 2002.

[10] J.B. Andresen et al., “Propagation Measurements and Models for Wireless Communications Channels,” *IEEE Communications Magazine*, Vol. 33, No. 1, January 1995.

[11] A. Chandrakasan, et al., "Power Aware Wireless Microsensor Systems", Keynote Paper ESSCIRC, Florence, Italy, 2002.

[12] K. Danilidis et al., “Real-time Tracking of Moving Objects with an Active Camera,” in *Real-Time Imaging Journal,* Vol. 4, No.1 pp.3-20. February 1998.