Comparison Between Various Application Independent Data Aggregation Techniques in Sensor Networks

Deepali Virmani[1] & Satbir Jain[2]

[1] Assistant Professor (IT Deptt.) , GPMCE, Delhi, India
[2]Assistant Professor (CSE Deptt.) , NSIT, Delhi, India

Summary

Sensor networks, a novel paradigm in distributed wireless communication technology, have been proposed for various applications including military surveillance and environmental monitoring. Such systems suffer bandwidth, energy, and throughput constraints that limit the quantity of information transferred from end-to-end. Mechanisms to perform data centric aggregation utilizing application specific knowledge provide a means to augmenting throughput, but have limitations due to their lack of adaptation and reliance on application specific decisions. We, therefore, propose three novel Application Independent data aggregation schemes :Signal Strength (aggregator with high signal strength), Fixed(pre-defined aggregators), Fly (adaptive) to perform application independent data aggregation in a time sensitive manner. Our work isolates aggregation decisions into a module that resides between the network and the data link layer and does not require any modifications to the currently existing MAC and network layer protocols. We take advantage of queuing delay and the broadcast nature of wireless communication to concatenate network units into an aggregate using a novel adaptive feedback scheme to schedule the delivery of this aggregate to the MAC layer for transmission. In our evaluation we show that bandwidth utilization has increased by as much as 80% , packet loss has been reduced by 55% in Fixed Application Independent Data Aggregation(FX-AIDA) , 65% in Signal Strength Application Independent Data Aggregation(SSB-AIDA) & 80% in Fly Application Independent Data Aggregation(FLY-AIDA) under heavy traffic loads. Average response time is reduced to a mark up difference Additionally, we show as much as average 70% reduction in transmission energy consumption with an overall reduction in header overhead . Time delay is reduced to 90% in case of mobility of the events by Fly Based Application Independent Data Aggregation. We also prove 95 % reduction in cost with FLY-AIDA that is major emphasis in sensor networks . We conclude our evaluation by proposing scheme Fly Based Application Independent Data Aggregation as best scheme through which sensor nodes can dynamically change from one aggregation technique to the other in an unpredictable environment and adapt to dynamic changes in the network and also support real time communication .

Key words : Data Aggregation ,SSB-AIDA,FX-AIDA,FLY-AIDA ,DOA

1. Introduction

The phenomenal growth in distributed wireless communication technology has led a novel paradigm known as sensor networks[2]. They have been proposed for use in various applications including military and civilian applications. Many dynamically changing scenarios such as battlefield, commercial inventory must be monitored using adaptive methods that utilize critical, real-time information gathered from integrated low powered sensors[9]. With large number of sensor devices being quickly and flexibly deployed in these networks, each sensor device must be autonomous and capable of organizing itself in the overall community of sensors to perform coordinated activities with global objectives. The sensors are programmed to listen for events. When an event occurs, the sensors inform the end point by generating wireless traffic[7]. As the number of nodes in the sensor network increases the probability of congestion near events increases. This localized congestion leads to degraded routing performance. Additionally, lot of packets get dropped and the over all response time increases. Further, sensors around the event spend considerable amount of energy to transmit packets which finally do not reach the end point. Data gathering is defined as the systematic collection of sensed data from multiple sensors to be eventually transmitted to the base station for processing. Since sensor nodes are energy constrained, it is inefficient for all the sensors to transmit the data directly to the base station. Data generated from neighboring sensors is often redundant and highly correlated. In addition, the amount of data generated in large sensor networks is usually enormous for the base station to process. Hence, we need methods for combining data into high quality information at the sensors or intermediate nodes which can reduce the number of packets transmitted to the base station resulting in conservation of energy and bandwidth. This can be accomplished by data aggregation. Data aggregation is defined as the process of aggregating the data from multiple sensors to eliminate redundant transmission and provide fused information to

Manuscript received July 5, 2008.
Manuscript revised July 20, 2008.

the base station. Data aggregation usually involves the fusion of data from multiple sensors at intermediate nodes and transmission of the aggregated data to the base station (sink)[1].We use the mechanism of AIDA[8] [10], an adaptive application independent data aggregation mechanism for sensor networks. AIDA performs lossless aggregation by concatenating network units into larger payloads that are sent to the MAC layer for transmission. Due to the highly dynamic and unpredictable nature of wireless communication in sensor networks, a novel feedback-based scheduling scheme is used to dynamically adapt to changing traffic patterns and congestion levels. By isolating our work in a layer that sits between the networking and data-link components of the communication stack, AIDA is able to perform such aggregation without incurring the costs of rewriting components to upper or lower layer protocols. We propose three application independent data aggregation techniques: Signal Strength based (aggregator with high signal strength),Fixed(pre-defined aggregators),Fly Based (adaptive) to perform application independent data aggregation in a time sensitive manner. The SSB-AIDA scheme attempts to identify the sensor which has the most useful information and the highest signal strength and assigns that sensor as the data aggregator to send packets to the end point. FX–AIDA scheme has the notion of predefined data aggregators in fixed regions of the sensor network region. Sensors surrounding the event send information to the aggregator which eventually sends only the most useful information to the end point. The mobility of events affects the performance of SSB-AIDA and FX-AIDA. We come up with the FLY-AIDA scheme which tries to combine the salient features from both the SSB-AIDA & FX-AIDA when we consider the mobility of the event. We have carried out a performance analysis of the three schemes with best and worst aggregation scenarios. Our analysis reveals that bandwidth utilization has increased by as much as 80% , packet loss has been reduced by 55% in FX-AIDA , 65% in SSB-AIDA & 80% in FLY-AIDA based scheme under heavy traffic loads. Average response time is decreased to a mark up difference Additionally, we show as much as a average 70% reduction in transmission energy consumption with an overall reduction in header overhead . Time delay is reduced to 90% in case of mobility of the events by FLY-AIDA . As well as we are able to achieve the goal of real time communication in sensor networks by the help of FLY-AIDA .The remaining part of the paper is organized as follows: In Section 2. deals with the design of the three application independent aggregation techniques, SSB-AIDA ,FX-AIDA and FLY-AIDA which takes advantage of the other two schemes. Section 3. describes the experiments performed and results obtained. Section 4 deals with the conclusion and the future work .

2. Three Application Independent Data Aggregation Techniques

2.1. No Aggregation

In No Aggregation scheme, sensor devices are unaware of other neighboring nodes. Each sensor upon detecting an event attempts to send the amount of information collected, however small it may be, to the end nodes (sink). Sensor devices do not apply any data aggregation technique and simply forward the data packets toward the sink node. As we can clearly see, such a scheme suffers from high packet dropping rate and low bandwidth utilization due to congestion in the network. Additionally, it also suffers from energy limitations as each device attempts to send packets received from multiple destinations irrespective of the importance of the data being transmitted[4]. Further more, the total amount of information received at the sink nodes would be less due to several packets getting dropped.

2.2. Signal Strength Based Application Independent Data Aggregation (SSB-AIDA)

In this scheme, the sensor network environment is divided into predefined set of regions. Each region is responsible for observing and reporting events that occur inside the region to the sink nodes. Also each sensor device inside the region sends data to neighboring sensor devices (only inside the region). Only one sensor, the **data aggregator**, sends the critical information received either from neighboring sensor devices or by itself to the sink node . A typical SSB-AIDA scheme is shown in Figure 1. As we see in the figure, all sensor devices inside the region detect the event. Each sensor transmits its signal strength to its neighbors. If the neighbor has a higher signal strength, the sender stops transmitting packets. After receiving packets from all the neighbors, the node that has the highest signal strength becomes the data aggregator and all other sensor devices stop detecting the event and helps only in routing the packet to the sink nodes. This scheme is highly suitable for environments where events have localized phenomenon, occurring in a fixed region of space.

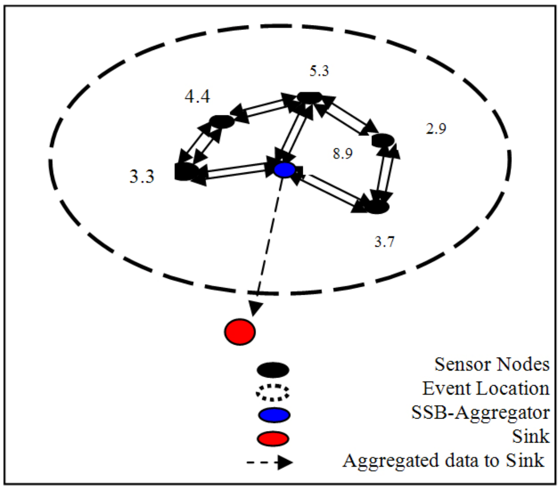

Fig. 1 SSB-AIDA Scheme

2.3 .Fixed Application Independent Data Aggregation(FX–AIDA)

As seen in the previous scheme, the sensor network environment is divided into predefined set of regions. Each region is responsible for observing and reporting events that occur inside the region to the sink nodes. In addition, in this scheme one sensor device based on geographical position with respect to either the sink or the center of the region is chosen as data aggregator. All other sensors inside the region are aware of this information. During event detection, all other sensors are supposed to send the event information to this data aggregator. The data aggregator after collecting data from other sensors sends only the critical information to the sink node. A typical FX–AIDA scheme is shown in Figure 2. As seen in the figure, during event detection, all sensors send data to the fixed aggregator. After collecting all data from other sensors, the aggregator sends only the critical information to the sink nodes. FX–AIDA adapts well to dynamic changes in the network topology and event mobility. If the event is highly mobile in nature, we see that many packets are exchanged between the sensors inside the region. But once the packets reach the aggregator, we see that only the most important information is sent to the sink nodes. Thus, FX–AIDA scheme reduces the traffic in such environments and makes sure the critical information is transmitted to the end nodes interested in the data. It also increases the throughput in such environments. However FX–AIDA scheme performs worse in environments where events are highly localized and mostly immobile in nature. We see that the data packets exchanged between the aggregator and other sensors inside the grid falls in the critical path. This increases the end-to-end response time. FX –AIDA scheme also increases congestion due to increased number of packets exchanged in the protocol compared to the SSB-AIDA.

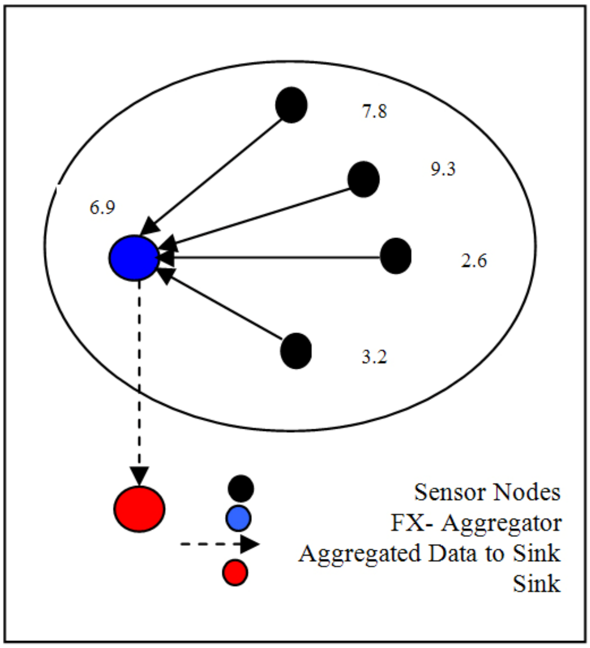

Fig. 2 FX-AIDA Scheme

2.4. FLY Based Application Independent Data Aggregation(FLY-AIDA)

Generally, the SSB-AIDA is preferred over FX-AIDA scheme in environments where events are highly localized. Due the advantages and disadvantages associated with each of the SSB-AIDA and FX-AIDA a hybrid approach of choosing schemes on the fly based on event duration and event mobility would be highly beneficial. Such an hybrid scheme would take the best of both the approaches. The basic approach of such a scheme is shown in Figure 3. As shown in the figure, every sensor initially is configured based on SSB-AIDA scheme. When a sensor detects an event, it first attempts to identify the sensor with the highest signal strength. In other words, the sensor which has the most critical and complete information about the event is identified. This is done the same way as described in the SSB-AIDA scheme. In addition each sensor also maintains a history of past events and the corresponding signal strengths the sensor detected. During event detection, each sensor checks its table for the previous entry and attempts to identify whether the event is highly mobile in nature or stationary. If it turns out that the event is localized, the SSB-AIDA scheme is followed and accordingly an aggregator is chosen. The event, it tries to send the information to the default fixed aggregator as described in FX-AIDA .

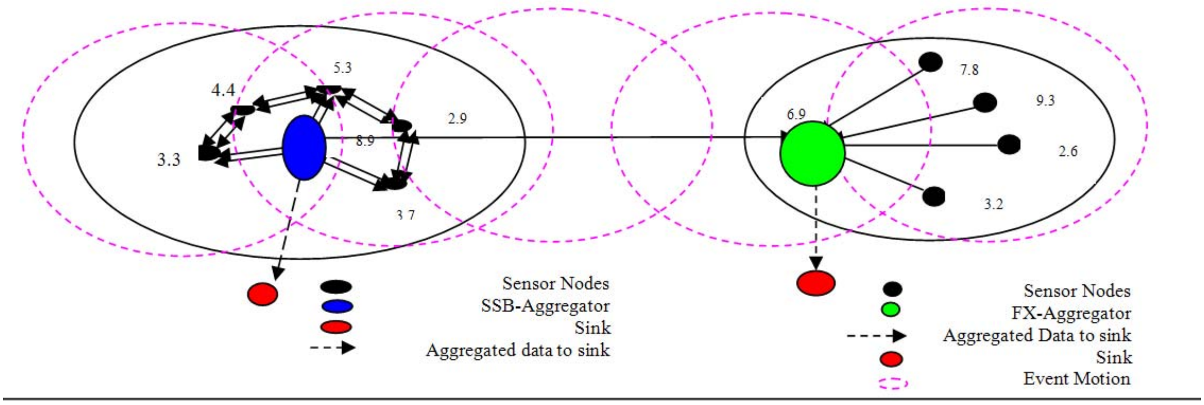

Fig. 3 FLY–AIDA Scheme

3. Experimental Results

In order to compare the different AIDA schemes discussed in the previous sections, we extended the functionality of the JSim[11] software package. Using this simulation framework we compared the AIDA techniques with the classic flooding (no aggregation) scheme. JSim is an event-driven simulator with extensive support for simulation for TCP, multicast protocols and also routing protocols in sensor networks. JSim supports different routing protocols; AODV, DSR [3],GPSR [4], etc. In this simulation, we fixed the routing protocol as DSR. To the JSim simulator , we added a CBR-Broadcast event for a CBR (Constant Bit Rate) traffic. We made use of this feature while simulating these AIDA techniques.

3.1. Simulation Testbed

For our experiments, we created a 100-node network. This network, which was randomly generated and was deployed over a 1000 x 1000 grid. The power of the sensor's radio transmitter is such that any node within a 100 meter radius is within the communication range and is called a neighbor node of a sensor. The radio speed (2 Mbps) and the power dissipation were set to default values. The processing delay for transmitting a message was chosen to be 5 ms. The size of each data packet was set to 200 bytes and the packet interval was set to 100 ms. Table 1 summarizes these network characteristics. Using this network configuration, we ran each AIDA scheme and tracked its progress in terms of rate of data dissemination, energy usage, through- put and average response time to reach the end nodes. The results of these experiments are presented in the following sections.

Table 1 : Characteristics for the 100 node wireless test network

Features	Values
Nodes	100
Grid	1000 * 1000
Radio Size	2 Mbps
Processing Delay	5 ms
Data Size	200 bytes
Data Interval Rate	100 ms

3.2 . Simulation Results

3.2.1. Average Response time

Figure 4 shows the comparison of the average response time for each of the aggregation schemes as compared with no aggregation scheme . Results show that with no aggregation scheme average response time increases as the number of the packet increases due to heavy localized congestion .With the SSB-AIDA as only the start up load is high to find the aggregator after the messages are passed smoothly so average response time is comparatively decreased to 60% in comparison with FX-AIDA where it is reduced to 50% .FLY-AIDA shows the most promising results as the average response time is reduced to 80%.

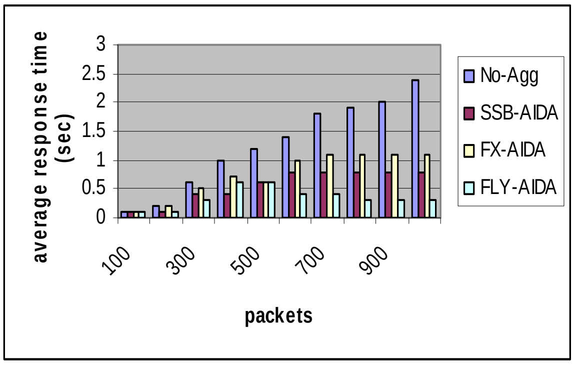

Fig. 4 Average Response time Comparison in various AIDA Techniques

3.2.2. Bandwidth Utilization

Figure 5 shows the throughput achieved by the network over time for each of the data aggregation proves the proper bandwidth utilization. As expected, No Aggregation scheme achieves very less throughput(bandwidth utilization) due to localized congestion. How ever the other three schemes achieves a considerably higher bandwidth utilization . Also, it is interesting to note that using the SSB-AIDA scheme, the system is able to achieve 65% bandwidth utilization which is comparable to the Perfect Aggregation scheme. This is due to the fact that the SSB-AIDA scheme has only the startup cost of finding an aggregator and the rest of the protocol remains the same as the Perfect Aggregation scheme. FX-AIDA shows 55% bandwidth utilization. But the simulation results prove that FLY-AIDA scheme proves to be the best in terms of bandwidth utilization as bandwidth utilization is increased to at most 80% as compared to no aggregation

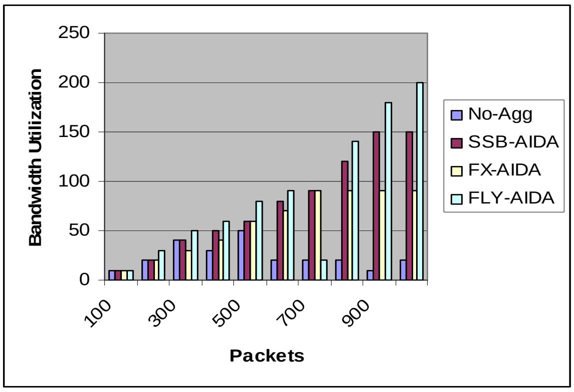

Fig. 5 Bandwidth Utilization Comparison in various Data AIDA Techniques

3.2.3. Energy consumed

Energy consumption, is adopted as another revealing metric to evaluate the performance of various AIDA schemes With limited power resources, it is vital for sensor nodes to minimize energy consumption during radio communication to extend the lifetime of the sensor network. AIDA achieves such energy savings via several approaches. First, AIDA reduces MAC channel contention costs by distributing these costs across multiple network units. Second, by using less MAC control packets, AIDA dampens congestion and reduces the number of collisions resulting in fewer retransmissions. In this experiment, we measure average energy consumption per delivered packet under 24 increasing traffic loads for three AIDA schemes. In Figure 6 energy metrics show that the scheme without data aggregation demonstrates the worst performance. SSB-AIDA shows the reduction in energy consumption of as much as 60% . FX-AIDA shows the reduction of 70% and the promising results are of FLY-AIDA the energy consumption is reduced to 80% as compared to no aggregation scheme . Thus by sacrificing a small, constant over-head in sending data only to the aggregator, three AIDA schemes achieves a dramatic reduction in system energy consumption .

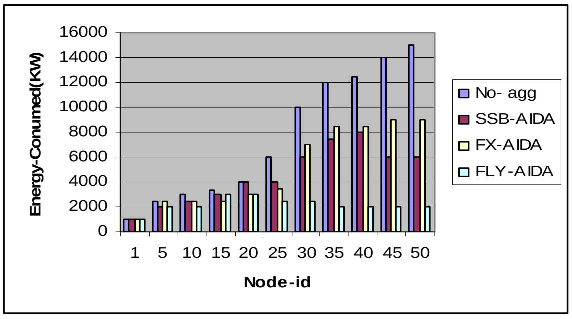

Fig. 6 Energy consumption Comparison in various AIDA Techniques

3.2.4. Packet Loss

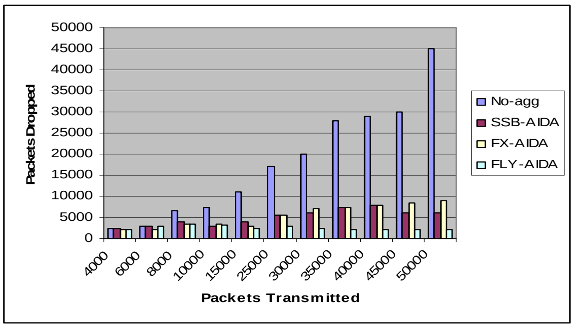

Fig. 7 Packet Loss Comparison in various Data AIDA Techniques

We also measured the loss rate for each sensor and the results are shown in Figure 7 . We see that the loss rate for the No Aggregation scheme is very high for some sensors. These are the sensors surrounding the event and the reason for huge packet loss is due to multiple re-transmissions as an effect of collisions in the network.

Since there is no data aggregation, all sensors are unaware of its neighbors signal strength and attempt to re-transmit packets in the event of collision. However we see that in the three schemes we proposed, the packet loss is considerably minimal due to data aggregation and reduced congestion. Packet loss has been dropped to 50% by FX-AIDA,65% by SSB-AIDA and most promisingly by FLY-AIDA Scheme to 80%.

3.2.5. Time Delay

$$D(K) = D_{min} + \# \text{Collisions}(K) * D_{resolve}. \quad (1)$$

D(K) → Delay experienced during time period [K, K+1]
D_{min} → Delay when no collision is experienced.
#Collisions(K) = No. of a collisions a successful transmission encounters at time interval [K, K+1]

$D_{resolve}$ = Sum of collision delay and time to resolve a single collision

Average collision probability p
DOA Degree of Aggregation
Average Number of transmission required for each successful transmission

$$E(\#Collision+1) = \frac{1}{(1-p)} \quad (2)$$

Expected no. of Collision each successful transmission will encounter.

$$E(\#Collision) = \frac{1}{(1-Z)^{N(K)/DOA(K)-1}} - 1 \qquad N/DOA \geq 1 \quad (3)$$

Relation between DOA values & MAC layer delay

$$D_{mac}^{(k)} = [D_{min} - D_{resolve}] + D_{resolve} (1-Z)^{\frac{1-N(K)}{DOA(K)}} \quad (4)$$

As in no aggregation scheme collisions are maximum so time wastage to achieve the results is maximum .The basic necessity of sensor networks is real time communication we are able to achieve this goal by FLY-AIDA as the time delay is minimized to only few m secs as shown in figure 8 .Simulation results prove 90% reduction in time delay with FLY-AIDA .

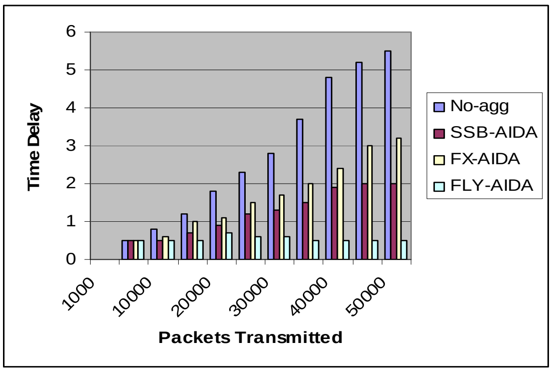

Fig. 8 Time Delay Comparison in various AIDA Techniques

3.2.6. Latency of events

One of the most important feature of sensor network is tracking mobile events[6]. Since the proposed FLY-AIDA scheme chooses the aggregation scheme on the fly, we define a series of terms in order to find the average latency. Let,

Total number of variations in mobility from low to high = V
Average latency of SSB-AIDA = L_{SSB}
Average latency of FX-AIDA = L_{FX}
Average latency of FLY-AIDA = L_{FLY}
Startup overhead of switching from one scheme to other = $O_{overhead}$

$$L_{FLY} = \frac{\min(L_{SSB}, L_{FX}) + O_{overhead} * V}{V} \quad (5)$$

Since the FLY-AIDA scheme, chooses the aggregation scheme on the fly based and adapts to the environment, the overall latency would be the sum of minimum of latencies of both the schemes and the startup overhead associated with switching from one data aggregation scheme to the other. This startup overhead occurs every time the Hybrid scheme changes from one scheme to the other. Hence the overall latency is calculated as given above . The results of this model are reported in Figure 9. As shown in this figure, we find that SSB-AIDA scheme performs well when the mobility of the event is low. As mobility increases, the SSB-AIDA latency increases exponentially due to more congestion. However, in the FX-AIDA scheme the latency increase is not exponential, thereby scales well with mobility of the event. In case of the FLY-AIDA scheme, we see that if the startup overhead of switching from one scheme to the other is less, the FLY-AIDA model performs to be the best .

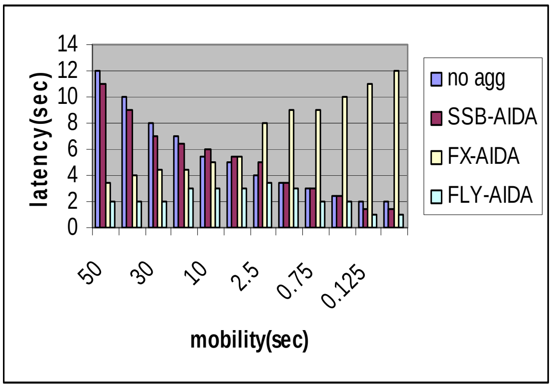

Fig. 9 Latency of events Comparison in various AIDA Techniques

3.2.7. Cost Saving

The cost of packet transmission in the simple single sender, single receiver scenario with no channel contention and an arbitrary MAC layer is the time consumed by the MAC acquiring and setting up each transmission plus the time for sending the message, all multiplied by the number of individual transmissions. Cost of packets transmission in sensor networks with aggregation is calculated as given below.

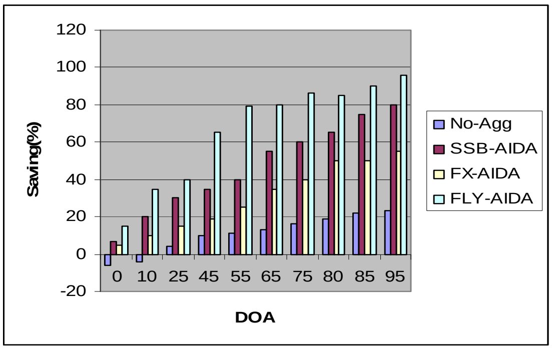

Fig. 10 Cost Saving Comparison in various AIDA Techniques

C_{agg} =M+(S*DOA+O) # R (6)

C_{none} = (M+S*R) * DOA (7)

C_{none} = Cost of sending packets without aggregation

$$\text{Saving percentage} = \left(\frac{\left(C_{NONE} - C_{agg}\right)}{C_{None}}\right) \quad (8)$$

$$= \left[1 - \frac{S*R}{M+S*R}\right] - \frac{M+O*R}{M+S*R} * \frac{1}{DOA} \quad (9)$$

Cagg → Cost of Sending packets with aggregation (Msec)

S → Network Size

R → Bytes/Seconds

O → Header overhead (bytes)

DOA → No. of packets aggregated

Cost (agg) → (m sec)

So as the more no of packets are aggregated the % saving in the cost increases .Experiment results prove that as the degree of aggregation increases FLY-AIDA shows the most promising results by reduction in cost to approximately 95%.

4. Conclusion & Future Work

The phenomenal growth in distributed wireless communication technology has led to the production of a wireless sensors which are capable of observing and reporting various real world phenomena in a time sensitive manner. However such systems suffer from bandwidth, energy and throughput constraints which limit the amount of information transferred from end-to-end. Data aggregation is a known technique addressed to alleviate these problems but are limited due to their lack of adaptation to dynamic changes in the network and unpredictable traffic patterns. In this paper we propose three application independent data aggregation techniques: SSB-AIDA (aggregator with high signal strength), FX-AIDA (predefined aggregators) to perform data aggregation. We come up with the FLY-AIDA (adaptive) scheme which tries to combine the features from both these two schemes when we consider the mobility of the event. Our simulation analysis reveals that average response time is reduced by 55% for SSB-AIDA ,65% for FX-AIDA and 80% for FLY-AIDA scheme as compared to the case without any data aggregation. Additionally, the bandwidth utilization is increased to at most 80% by FLY-AIDA . Time delay and packet loss is reduced to a markup difference .Energy consumption a very important point in sensor networks is reduced to 50% by FX-AIDA ,60% by SSB-AIDA and 80% by FLY-AIDA .We are able to achieve to real time communication by help of FLY-AIDA by reducing the time delay. Cost in sensor networks with data aggregation is reduced a lot. .Reduction of 95% of cost is achieved by FLY-AIDA We conclude our evaluation by proposing an analytical model for FLY-AIDA scheme through which sensor nodes can dynamically change from one aggregation technique to the other in an unpredictable environment and adapt to dynamic changes in the network. In short, data aggregation hold the promise of achieving high performance at a low cost in terms of complexity, energy, computation and communication. Although our initial work and results are promising, there is still a great deal of work to be done in this area. Different scenarios for the wireless network with more sparse sensors deployed needs to be tested. The adaptive Fly-AIDA scheme is only proposed but needs to be implemented under various scenarios. We also need to consider the impact of multiple events in data aggregation. Transactions on Embedded Computing System Special issues on Dynamically Adaptable Embedded Systems.

5. References

[1] C. Intanagonwiwat, D. Estrin, R. Govindan, and J. Heidemann. "Impact of network density on data aggregation in wireless sensor networks. Technical Report ", University of Southern California, 2005

[2] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion " A scalable and robust communication paradigm for sensor networks". In MobiCom, 2004.

[3] D. B. Johnson and D. A. Maltz. "Dynamic source routing in ad hoc wireless networks". In Mobile Computing, 200

[4] Karp and H. T. Kung. GPSR "Greedy perimeter stateless routing for wireless networks". In MobiCom, 2005.

[5] C.Y. Wan, S. Eisenman, and Andrew Campbell. CODA: " Congestion detection and avoidance in sensor networks". In SenSys 2003 .

[6] E. M. Royer and C. E. Perkins. "An implementation study of the AODV routing protocol ". In Proceedings of the IEEE Wireless Communications and Networking Conference, 2006.

[7] G.J. Pottie, W.J. Kaiser, "Wireless Integrated Network Sensors,"*Communications of the ACM*, vol. 43, no. 5, pp. 551-8, May 2003.

[8] Krishnamachari B, Estrin D and Wicker S 2002, "The Impact of Data Aggregation in Wireless Sensor Networks", Proc. of the 22nd Int. Conf. on Distributed Computing Systems ,2004.

[9] Khemapech, Duncan I and Miller A, 2005, "A survey of wireless sensor networks technology", in PGNET, Proceedings of the 6th Annual PostGraduate Symposium on the Convergence of Telecommunications, Networking & Broadcasting, EPSRC.

[10] T. He, B.M. Blum, J.A. Stankovic, and T.Abdelzaher, "AIDA: Adaptive application independent data aggregation in wireless sensor networks," *ACM Transactions on Embedded Computing Systems*, vol. 3, no .2, May 2004, pp. 426-457.

[11] http/www/jsim.org

Deepali Virmani Received the B.E. Degree from Maharishi Dayanand University, Rohtak (Haryana) in 2001 & M.Tech Degree From Guru Gobind Singh Indraprastha University, Delhi in 2005. Presently she is carrying out the research work in the area of sensor network for the **Doctorate Degree** from Delhi University. After working as a lecturer in Engineering College, she has been an Asstt. Professor in department of Information Technology at Indraprastha University since 2007. Her Research Interest includes wireless sensor network and adhoc network. She is a member of IEEE, IETE & ISTE India.