ABSTRACT
In all types of communication network, congestion is always an important issue. Congestion in wireless network cause loss of packets, longer delay and throughput degradation. In applications like multimedia, nodes typically produce large amount of data necessitates high transmission rates and more processing, a fundamental concern in Wireless Multimedia Sensor Networks. Many sources send data of different applications so weighted fairness is also required. To handle this type of scenario priority based congestion control techniques were discussed. The aim of this paper is to survey existing priority based congestion control techniques to classify them & to discuss their advantages and drawbacks.

General Terms
Priority based congestion control algorithms.

Keywords
Congestion control, Wireless Sensor Networks, QoS.

1. INTRODUCTION
The internet and wireless technologies are growing rapidly and have been achieved great success in the last few years, but due to increase in network size and no. of users, network congestion occurs which causes longer delay, packet loss and other performance factor degradation. Computing is moving to an era where applications require lots of task performance. Such applications include multimedia applications, high frequency sensing applications, file transfer etc. If devices, enabling these applications are going to become an integral part of tomorrow’s networks, solutions to reduce the effects of congestion in wireless networks are required.

The Wireless Sensor Network has gained plenty of interest in both academia and industry in recent years [1]. Due to merging nature of upstream traffic, congestion most likely appears in the upstream direction. Because of the congestion consequences, it must be efficiently controlled.

In WSNs, Congestion may be distinguishing between two different types [2]:

a) Node Level congestion
b) Link level congestion

Node level congestion: It is common in typical multi-hop networks. It occurs in situations where the buffer of intermediate sensor overflows. Node level congestion results in an increase of both packet loss and delivery time, since packets have to be waiting within the queue of intermediate nodes before they are forwarded to either another node or a sink. Additionally, nodes consume additional energy due to the obligatory retransmission of rejected packets.

Link Level congestion: It is usually happened because of collisions done by concurrent transmissions of neighboring nodes. Link level congestion can result in increased packet loss due, increase packet delivery time and decreased channel utilization and overall throughput. Moreover, packet retransmission is too necessary, consequential to network energy dissipate.

Congestion control normally follows three steps:
1) Congestion detection
2) Congestion Notification
3) Rate adjustment.

Congestion Detection’s main mechanisms are:

a) Queue occupancy: If a node receives more data than it can forward, the additional data have to be stored in the limited buffer space [2]. In order to avoid buffer overflow each node in the network should monitor its residual buffer space and when it exceeds a predefined threshold, then the node has to inform its upstream neighbors. Then the upstream neighbors have to reduce their sending rate towards that node and even better to avoid using that node as an intermediate node. If the neighboring nodes are not intimated on time, then the buffer of the specific node overflows and the occurrence of congestion begin.

b) Channel utilization: Nodes that use the CSMA protocol can quite easily listen to the channel, and in case that they sense the channel busy they have to calculate the local channel utilization degree [2]. If a sensor senses that the channel loading reaches a certain fraction of the channel capacity it has to postpone its data transmission in order to avoid packet collisions.

Congestion Notifications [3] are of two types:

a) Explicit Notification (ECN): It is a particular control packet that notifies the congestion by the congested nodes, to its neighboring nodes.

b) Implicit Notification (ICN): In this type of notification, congestion warning is piggybacked by the congested node in the header of regular data packets.

Congestion Avoidance [3] basic mechanism’s are:

1) Rate adjustment: Rate adjustment refers to modifying the transmission rate of sending nodes upon the reception of the congestion notification.

2) Traffic Redirect: In traffic redirection, the nodes dynamically assign its leaving traffic to the uncongested route. The congested paths are avoided via the feedback information acquired from the neighbor such as the high loss rate of the links.

The remaining paper is organized as follows: Section 2 describes the procedures for the congestion control unit for priority based congestion control. Section 3 Explains
Comparison and Analysis. Lastly, the manuscript is concluded in section 4.

2. PRIORITY BASED CONGESTION CONTROL

2.1 Priority based Congestion Control in Wireless Sensor Networks (PCCP)

PCCP suggests intelligent congestion detection (ICD) that detects congestion by mean packet inter-arrival time and mean packet service times on the MAC layer. ICD describes a new congestion index, congestion degree $d(i)$, which is defined as the ratio of the average packet service time over average packet inter-arrival time in a pre-specified time interval in each sensor node.

Proposed work utilizes ICN in every node that used to piggyback congestion status in the header of regular data packets that has to be transmitted.

The piggybacked information at a sensor node $i$ include mean packet service time ($t_{s_i}$), mean packet inter-arrival time ($t_{a_i}$), global priority $GP(i)$, and the number of offspring node $O(i)$. A node, then computes its global priority index by summating its source traffic priority index and all the global priority index of its child nodes, which is piggybacked in the received data packets. Proposed protocol uses the Priority based Rate Adjustment (PRA).

It introduces a scheduler by two sub queues connecting the network layer and the MAC layer. If the scheduling rate $r_{svc}^i$ is kept below the MAC forwarding rate $r_f^i$, the output rate will approximately equal the output rate $r_{out}^i$. Therefore, through adjusting the scheduling rate $r_{svc}^i$, congestion could be avoided or mitigated.

Fig. 2 Node model in PCCP

Congestion degree $d(i)$ and priority index ($TP(i)$ and $GP(i)$) introduced here provides more information than the CN bit and enables exact rate adjustment. PRA adjust the scheduling rate $r_{svc}^i$ and the source rate $r_{src}^i$ at each sensor node after overhearing congestion notification from its parent node, in order to control both link-level congestion and node-level congestion.

1) It attains high link utilization & flexible fairness.

2) It manages the small buffer size. Consequently, it can avoid/reduce packet loss & therefore improve energy-efficiency & provide longer delay.

But it does not take current queue utilization into account, which leads to bigger queuing delays & frequent buffer overflows accompanied by increased retransmissions. It also doesn’t have a mechanism for handling prioritized heterogeneous traffic.

2.2 A QoS Adaptive Congestion Control in Wireless Sensor Network

This paper [4] proposed a new QoS adaptive cross-layer approach to control the congestion and maintain QoS guarantee for different application data on sensor network. This approach classifies congestion control algorithm as namely short-term and long-term congestion. To make sure real time and non-real time data stream, hop-by-hop QoS conscious scheduling and QoS distributed MAC Manager are considered.

This algorithm used Congestion Scale for congestion detection. It is defined by the ratio of the mean packet service rate and mean packet inter-arrival rate.

When congestion happens implicit congestion notification is used and avoid control message.

Node Model is shown in Fig. 3 there are two active queues, one for real time and other for non real time applications and one backup queue for storing unacknowledged non-real time data. Classifier classifies the packet, and scheduler schedules the queue according the application priority.

For Congestion mitigation they proposed two congestion control mechanism Long-term congestion control and Short term congestion control.
Moreover, the different priority application is used for both real time and non-real time data using node priority depending on weight factor.

Implicit piggyback message carrying information of different parameter assists the upstream node in order to get proper measures to avoid congestion. Extensive simulations are completed and it shows improved performance in terms of energy loss and average queue occupancy than two other existing schemes CCF(Congestion Control and fairness) [5] and NCC(No congestion control).

2.3 PHTCCP (Prioritized Heterogeneous Traffic oriented Congestion Control Protocol)

This protocol[13] makes sure proficient rate control for prioritized heterogeneous traffic. It uses intra-queue and inter-queue priorities for making certain feasible transmission rates of heterogeneous traffic.

PHTCCP uses packet service ratio \( r(i) \) to measure the congestion status at each node. Packet service ratio is described as the proportion of average packet service rate \( (R_s') \) and packet scheduling rate \( (R_{sch}) \) in every sensor node \( i \) which is

\[
   r(i)=\frac{R_s'}{R_{sch}^i}
\]

At this point, the packet service rate \( R_{sch} \) is the inverse of packet service time \( t_s' \). The packet service time \( t_{si} \) is the time gap when a packet arrives at the MAC layer and when it is successfully transmitted towards the next hop. \( t_{si} \) includes packet waiting time, collision resolution, and packet transmission time at MAC layer. In order to obtain \( R_s \), the average packet service time, \( t_{si} \) is calculated using exponential weighted moving average formula (EWMA). By using EWMA, \( t_{si} \) is changed every time a packet is forwarded as,

\[
   t_{si}'= (1-w_s) \cdot t_{si}' + w_s \cdot inst(t_{si}')
\]

While \( inst(t_{si}') \) is the instantaneous service time of the packet that has just been sent and \( w_s \) is a constant value while \( 0<w_s<1 \).

The packet service ratio shows the congestion level at every node. When the ratio is come to 1, the scheduling rate is identical to the forwarding rate, when the ratio is bigger than 1, the scheduling rate is less than the average packet service rate. Both of these cases points out the decrease of the level of congestion. When it is less than 1, queuing up of packets happens at the queue. This too indicates link level collisions. Thus, the packet service ratio is an efficient measure to identify both node level and link level congestion.

PHTCCP uses the Implicit Congestion Notification. Each node \( i \) uses piggybacking to send its packet scheduling rate \( R_{sch}^i \).

\[
   \text{Total number of child nodes,$C(i)$ .}
\]

Number of active child nodes at time \( t \), \( A(C(i)) \) and

Weighted average queue length of its active child nodes in its packet header. Because of the broadcast transmission of the wireless channel, all the neighbor nodes, including child nodes of the node \( i \) overhear the congestion notification value. When the value of \( r(i) \) goes beneath an assured threshold (according to the application necessity), the rate adjustment procedure is activated.
PHTCCP uses hop-by-hop rate adjustment which ensures that heterogeneous data reach to the base station at their desired rates. The output rate of a node is regulated by changing the scheduling rate, $R_{sch}^i$.

The change in rate depends on two cases:

When node $i$ determines that all the child nodes of its parent (including itself) are active, (in Figure 4(a))

At $(C_{p_j}) = C(p_i)$, then node $i$ makes adjustment in its scheduling rate. In this case, every node $i$ put its scheduling rate equal to $1/C(p_i)$ th of its parent’s scheduling rate. In Figure 4(a), if the scheduling rate of the parent node is $R_{sch}^{p_j}$, each child node has the scheduling rate, $R_{sch}^{p_j}/4$. This ensures that the total scheduling rate of all the child nodes is not greater than the scheduling rate of their parent node.

When node $i$ finds out that some of the child nodes of its parent (i.e., its siblings) are idle that is when $At(C(p_i)) < C(p_i)$, it again adjusts its scheduling rate. So that available link capacity will be distributed to active child nodes.

PHTCCP is energy efficient and provides lower delay. It is also feasible in terms of memory requirement considering the configurations of today’s multipurpose (temperature, humidity, light) mote.

2.4 Priority based rate control for service differentiation and congestion control in multimedia sensor networks.

Yaghmaee et al.[6] presents a priority based rate control method for congestion control and service differentiation in Wireless Multimedia Sensor Networks. The proposed congestion control technique can regulate the source traffic rates based on current congestion status in the upstream nodes and the priority of every traffic source. The proposed model is able to support four dissimilar traffic classes that are:

real traffic class (EF class), high priority non real-time traffic class (NRT1 class), medium priority, non real-time traffic class (NRT2 class), and low priority non real-time traffic class (NRT3 class).

A congestion control unit used in paper made up of two key units that are Congestion Control Unit (CCU) and service differentiation unit (SDU).

A congestion control unit made up of three parts that is congestion detection unit (CDU), congestion notification unit (CNU), and rate adjustment unit (RAU).

CDU evaluates the congestion detection by taking the difference between its input rate and maximum allowable output rate.

RAU computes the new rate for each child traffic source as well as its local traffic source, based on a recent congestion index and the source traffic priority. The latest rate is sent to CNU component which is responsible for informing, by adding the new rate of each child node of the sending data of each node.CNU utilizes implicit congestion notification by inserting the new rate of each node in the sending data of each node. Once a node detects congestion, it is likely to regulate traffic rate appropriately.

The service differentiation unit used for classifying the network traffic for facilitating the QoS in the network. The proposed models keep four different traffic classes specially:

Real time traffic class (EF class),
High priority, non real-time traffic (NRT1 class),
Medium priority, non real-time traffic (NRT2 class),
Low priority, non-real time traffic (NRT3 class).

They make use of separate queue for every traffic class. For non real time classes RED protocol is used which assist to obtain service differentiation in nodes. In proposed algorithm,
it defined overall input rate and overall output rate. Each parent node i compute its total input rate from all of its child nodes. Moreover, each parent node i computes and propagates the new output rate of each of its child nodes.

Based on algorithm when some child nodes do not have sufficient traffic, the input rate to the parent node will be decreased, that will in turn lead to an increase in sending rate of other child nodes. The amount of the raise in the sending rate is not same for all child nodes. It depends on the global priority of the child nodes. The higher priority node will get a much increment in sending rate than a low priority mode, while some child nodes send enough traffic, and afterward input rate to the parent node i will turn into greater than its previous value. This creates the value of $\Delta r_i$ to become negative. So to prevent any packet loss or more delay, every child node reduces its sending rate. The amount of reduction is dependent on the global priority of each node to assure in fairness and in elevated bandwidth utilization.

This paper evaluated the performance of the proposed model in different cases and with different scheduling protocols WRR (Weighted Round Robin) and PQ (Priority Queuing). It is found that it is likely to provide low queuing delay and assured bandwidth for high priority real time traffic. It also achieved less packet loss probability, with their proposed congestion control algorithm.

2.5 Enhanced Congestion Detection and Avoidance for Multiple Class of Traffic in Sensor Networks.

This paper [7] proposes energy proficient congestion control method for sensor networks named ECODA, which include three mechanisms:

a) Utilize dual buffer thresholds and weighted buffer difference intended for congestion detection.

In this method Buffer is described as three states “accept states”, “filter state” and “reject state” as revealed in Fig.7. Two thresholds $Q_{min}$ and $Q_{max}$ be there to edge dissimilar buffer states.

b) Flexible Queue Scheduler for packets scheduling

In Queue model, there are two sub-queues. First is intended for local generated traffic, and the second is intended for transit traffic. Transit traffic queue contains the packets which are collected from different sources. For each source, packets are sorted through their dynamic priority as of high to low. When transmitting next packet, a round robin algorithm is executed. To make certain fairness, the algorithm scans the transit traffic queue starting head to tail. One packet from individual source is transmitted from the transit traffic queue, and after that a local produced packet is transmitted.

c) A bottleneck node based source sending rate control scheme.

1) Find out the routing path status of a certain node to sink.

A node for whose next node is sink, its data forwarding delay is piggybacked in data packets. Its child overhears this value and match up to its value with parent value and performs this computation:

$$D_{max}(i) = \text{MAX} \{ D(p), D(i) \}$$

Where $D(p)$ is delay of parents, $D(i)$ is delay of nodes.

This practice is recursively calculated to source node.

2) Bottleneck node finding and source data sending rate control.

Using that scheme, source data sending rate can be keeping pace more precisely. Upon receiving a backpressure message, the source node or forward node decreases its data sending rate or adjusts the data transmission rate for different paths if multiple paths are present. However, dissimilar existing protocols, if no backpressure message is received, the source node doesn’t raise its data sending rate additively. In ECODA, the source data transmission rate can be precisely set as $1/D_{max}$ upon receiving a message from downstream node.
ECODA manage transient congestion and persistent congestion effectively. For transient congestion, it takes up hop-by-hop congestion control method. Intended for persistent congestion, it gets up bottleneck node based source data sending rate control. ECODA protocol attains high bandwidth utilization and flexible fairness. It reduces packet loss, improve energy efficiency, lesser delay and more throughput than CODA [8].

2.6 Priority Based Congestion Control For Heterogeneous Traffic in Multipath Wireless Sensor Networks

This paper [9] proposed a congestion control technique for heterogeneous traffic in the multipath wireless sensor network. The proposed protocol assigns bandwidth relative to the priority of a lot of applications concurrently running in the sensor nodes. Every node route its locally produced data as well as the data produced from other nodes. The parent node of every one sensor node assigns the bandwidth based on the source traffic priority and transit traffic priority of the data from varied or heterogeneous application in the child nodes. Congestion detection is done by the packet service ratio and congestion notification is implicit. Rate adjustment is done by hop-to-hop. It makes sure that the heterogeneous data from different stations will reach the base station at the preferred rates.

Fig.9 Bottleneck node based source data sending rate control.

ECODA introduces the bottleneck node based source data sending rate control. ECODA protocol attains high bandwidth utilization and flexible fairness. It reduces packet loss, improve energy efficiency, lesser delay and more throughput than CODA [8].

2.7 Priority Node based Congestion Control Algorithm For Wireless Sensor Networks

Sun et al.[10] proposed an algorithm which substantiates prioritization according to the degree of congestion state, and after that adaptively decides its probability of accessing channel in order to rapidly ease the congestion in the area. The essential idea of PNBCC is to classify node priority by the queue load when a node has limited queue buffer, consequently select different congestion control strategy. The proposed algorithm provides two thresholds. In additions, as in the 802.11 distributed coordination function (DCF) mechanism, back off time reflects the capability of a node’s access to the channel. The small back off time is, the more contention ability of the sensor node is. This paper uses improved DCF in MAC layer. A modified algorithm place ahead an improved back off algorithm combined with the node priority on the basis of DCF. Priority factors reflect the ability of node accessing the channel. The higher the priority of node, channel access to the node is easier. In Back off algorithm of 802.11, in order to effectively avoid collision which occurs frequently in the network, most algorithms adopt a binary exponential back off algorithm.

After successfully sending the data, the node’s contention window returns back to its minimum value, which is helpful to the next channel contention, but reduces other node’s probability of channel contention. This phenomenon is called capture effects, which cannot make sure fairness of nodes occupying the channel. For this consideration, this paper employs modified DCF in the back off stage. In this enhanced algorithm, the contention window will remain same after each successful transmission. Only after c successive successful transmissions the window is halved. An algorithm is presented as below:

$$CW[AC]_{\text{max}} = \begin{cases} \frac{1}{2} \cdot CW[AC]_{\text{max}} \cdot n \geq c \\ CW[AC]_{\text{max}} \end{cases}$$ (1)

Else collisions occur,

$$CW[AC]_{\text{max}} = \min(CW[AC]_{\text{max}} \cdot 2 \cdot CW[AC]_{\text{max}}).$$ (2)

Where $n \in (0, m)$ is called “back off stages” and $m$ is the maximum back off stage.

PNBCC can rapidly alleviate the congestion situation around the congested nodes. In the meantime, using flexible contention mechanisms greatly decrease probability of collision amongst nodes, which also get better network performance. Also

PNBCC’s packet losses are lesser than other algorithms FUSION [11], CODA [8], HCC [12].

3. COMPARISON AND ANALYSIS

From the discussion of above priority based congestion control schemes of WSN it can be seen that they differ in

5. REFERENCES


Table 1: Comparison of Priority based Congestion Control Techniques

<table>
<thead>
<tr>
<th>Protocols</th>
<th>Congestion Detection</th>
<th>Congestion Notification</th>
<th>Congestion Avoidance</th>
<th>Techniques Introduced</th>
<th>Performance metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCCP[1]</td>
<td>Intelligent Congestion Detection (ICD)</td>
<td>Implicit</td>
<td>Rate adjustment (Exact)</td>
<td>Node priority introduced</td>
<td>Normalized throughput, Queue length and weighted fairness</td>
</tr>
<tr>
<td>Quality of Service Adaptive Congestion Control[4]</td>
<td>Congestion scale</td>
<td>Implicit</td>
<td>Rate adjustment, alternate path</td>
<td>QoS adaptive cross layer approach</td>
<td>Average Queue occupancy, Average residual energy</td>
</tr>
<tr>
<td>PHTCCP[13]</td>
<td>Packet service ratio</td>
<td>Implicit</td>
<td>Rate adjustment (Exact)</td>
<td>Inter-queue and intra-queue priorities</td>
<td>Normalized throughput, delay Energy Efficiency</td>
</tr>
<tr>
<td>Priority based rate control for service differentiation[6]</td>
<td>Congestion index</td>
<td>Implicit</td>
<td>Rate adjustment (Exact)</td>
<td>Distinguish and service real time traffic from non real time traffic accordingly</td>
<td>Normalized throughput, loss probability, delay</td>
</tr>
<tr>
<td>ECODA[7]</td>
<td>Buffer, Weighted buffer difference</td>
<td>Implicit</td>
<td>Rate adjustment (Exact),multiple path</td>
<td>Dynamic priority used</td>
<td>Throughput ,delay and buffer occupancy</td>
</tr>
<tr>
<td>Priority based congestion control for multipath[9]</td>
<td>Packet service ratio</td>
<td>Implicit</td>
<td>Rate adjustment (Exact),multiple path</td>
<td>Used For multipath scenario</td>
<td>Normalized throughput using fairness</td>
</tr>
<tr>
<td>PNBC[10]</td>
<td>Ratio of current buffer length and current queue length</td>
<td>Implicit</td>
<td>Rate adjustment (Exact)</td>
<td>Classify node priority by the queue length</td>
<td>Throughput, packet loss, channel occupancy time, Average delay time</td>
</tr>
</tbody>
</table>

terms of congestion detection, congestion notification and avoidance policy collectively. These schemes use different metrics to compare and find their algorithm better than existed algorithm. As illustrated in Table 1, it can be seen that each scheme introduced some techniques for improvement so that its results are better in terms of the performance metrics like throughput, delay, fairness etc., when compared to other previous existed congestion control schemes.

4. CONCLUSION AND FUTURE WORK

The ad-hoc network has attracted attention intending at the achievement of everywhere in the world. Applications of adhoc network are communication between vehicles and sensor network and so on. Amongst them, wireless sensor network is useful in an emergency like disasters, medical, etc. In this application, one of the troubles is congestion for increasing of traffic of the urgent and medical data packets on wireless communication paths. It is a problem to be solved, to guarantee quality of service in ad-hoc network in an urgent situation. In this paper, we took review on priority based congestion control in wireless sensor Network. Based on literature survey it is found that enough Congestion Control technique is for homogeneous environment & for Congestion control. In a Heterogeneous traffic environment like multimedia traffic and emergency it is still an issue of concern. In real time traffic if packets are delayed up to some extent then they are useless so if more parameters could be added for finding the Dynamic Priority of the packet and so of nodes then congestion could be avoided more efficiently and accurately.

...


