Cross Layer Design for Service Differentiation in Mobile Ad Hoc Networks

Zhongbang Yao, Pingyi Fan, Zhigang Cao
Department of Electronic Engineering
Tsinghua University
Beijing, China
zbyao@eee.hku.hk

Victor O. K. Li
Department of Electrical and Electronic Engineering
The University of Hong Kong
Hong Kong, China
vli@eee.hku.hk

Abstract—Cross layer design is a promising approach in mobile ad hoc networks (MANET) to combat the fast time-varying characteristics of wireless links, network topology, and application traffic. In this paper, we employ cross layer design to develop a novel scheduling scheme with two optimizations aimed at service differentiation. The scheduling scheme is executed at the network layer of every station according to the channel conditions estimated by the MAC layer. The optimizations are based on traffic property sharing and packet timeout period interaction to reduce the packet collisions and improve network performance. We evaluate the proposed scheme under different network loads in terms of packet delivery ratio, average end-to-end delay and delay jitter. The simulation results show that our scheme can provide different service differentiations for time-bounded and best effort traffics. In particular, we can guarantee the delay and delay jitter requirements of time-bounded traffic.

I. INTRODUCTION

A mobile ad hoc network (MANET) consists of mobile stations, which communicate with each other, without any infrastructure, such as the base stations, as in cellular networks. Much prior research has shown that application traffic and wireless characteristics affect MANET design in each of the seven layers in the ISO/OSI reference model. Intuitively, through information sharing and interaction among different layers, one may take more efficient actions and respond more quickly to the time-varying changes of the network. This is the cross layer design method.

Some research has been done in cross layer design for ad hoc networks [2,3]. In [2], channel reservation control packets are used to estimate the channel condition, which is used by media access control (MAC) and routing to implement rate adaptation and optimal route selection. Reference [3] proposes a new multiple access control scheme composed of two phases: scheduling and power control, both based on the physical layer parameters, such as signal-to-interference-noise-ratio (SINR), etc. But this scheme assumes that there exists a separate feed back channel to send SINR measurements and a central controller to execute the scheduling algorithms, which may not be available in practical ad hoc networks.

There is also much work in scheduling and priority provision in wireless networks [4-10]. Chun et al. [4] evaluate the performance of different packet scheduling algorithms in the network layer. In addition, they also investigate the queuing dynamics at the mobile stations. In [5], three mechanisms to support message priorities in the MAC layer are presented: packets with different priorities stored in different FIFOs, assigned different access deferral and different backoff times. Reference [6] also proposes three different service differentiation schemes for IEEE 802.11 MAC layer – using different contention windows, assigning different inter frame spacings (IFSs), and using different maximum frame lengths for different users. However, only the network or MAC layer behavior is considered in [4-6], while the impacts of other layers are ignored. In [7], a busy tone priority-scheduling (BTPS) scheme is presented, in which two narrow-band busy tone signals are used to guarantee high priority stations preferred access to the wireless channel. But this scheme needs to divide the wireless spectrum into three channels: BT1, BT2 and Data channels. H. Luo, S. Lu and V. Bharghavan [8] propose a new model to address the trade-off between fairness and channel utilization, which concentrates on ensuring the fair allocation of channel bandwidth and maximizing the spatial reuse. Some other scheduling and priority algorithms are also proposed in wireless cellular networks or wireless LANs, such as [9, 10].

Generally, the proposed scheduling and service differentiation schemes in wireless networks are mainly aimed at cellular or wireless LAN networks. The schemes presented in ad hoc networks mostly concentrate on MAC or network layer, and few schemes consider the interaction between layers.

In this paper, we employ cross layer design to develop a scheduling scheme and two optimization methods to provide service differentiation in MANETs. In our scheme, we estimate the channel condition in the MAC layer and send this information to the network layer, enabling it to take different actions for priority queuing. Considering the scalability problem of flow-based scheduling algorithms, which are mostly employed by the scheduling algorithms in wireless LAN or cellular networks, we adopt the class-based scheduling scheme. To improve the traffic transmission quality at the receivers, two optimizations are proposed through traffic properties sharing and packet timeout period interaction among different layers.

The rest of this paper is organized as follows. Section II presents the scheduling and optimization scheme for service differentiation through cross layer design. Section III evaluates

the proposed scheme and shows the simulation results. We give our conclusions in Section IV.

II. SCHEDULING AND OPTIMIZATION FOR SERVICE DIFFERENTIATION THROUGH CROSS LAYER DESIGN

In this paper, we consider two traffic classes: best-effort traffic and time-bounded traffic (e.g., voice, video). When a packet arrives from the application layer at the network layer, it is put into different fist-in-first-out (FIFO) queues. Based on the different channel conditions estimated by the MAC layer, the network layer takes different actions when reading packets from the queues. At the same time, the MAC layer takes different actions according to different traffic types and network packet timeout requirements.

A. Scheduling and Optimization Model through Cross Layer Design

Fig. 1 shows the framework of our proposed scheduling and optimization scheme. Since we do not consider the impact of the transport layer, and just use simple transport protocol, such as UDP, the transport layer is ignored in Fig.1.
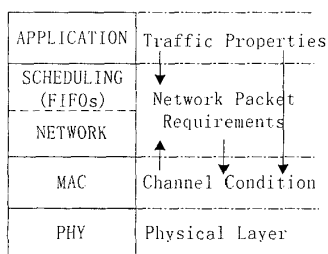

Fig. 1. Scheduling and optimization through cross layer design

From Fig.1, information exchanged between layers includes: traffic properties, network packet requirements, and channel conditions.

Traffic property information is shared among the application, network, and MAC layers. In our model, the traffic properties include: traffic type, packet transmission delay bound, and packet transmission delay jitter bound. Traffic type is set as the type-of-service (TOS) field and traffic category (TC) field in the network and MAC layers, respectively. Delay and delay jitter bounds are pre-defined values set by the application layer that are conveyed to the network and MAC layers through inter-layer interfaces. The application layer can determine whether to receive a packet based on the traffic properties after the packet arrives from the network layer.

In each station, after transmitting a packet, the network layer must wait for the MAC layer to inform it the packet is successful or not, or for the transmission timeout. If MAC informs the network layer the transmission outcome before the packet timeout, the network layer can initiate the re-transmission process, or just discard the packet and transmit the next one. However, if the packet timeout occurs first, the network layer will send the same or another packet to MAC when it is still transmitting the previous one, which will cause useless transmissions and severe network contention. Thus, to avoid MAC transmitting the packets that have already timed out at the network layer, packet requirements are conveyed to the MAC layer. Here, the requirements mainly refer to the packet timeout periods for the best-effort and time-bounded traffics denoted by T_{rb} and T_{rv}, respectively. Therefore, based on the timeout periods, traffic delay bounds and traffic delay jitter bounds, the MAC layer can determine whether to transmit or not after reading a packet from its FIFO.

The estimated channel condition at the MAC layer is conveyed to the network layer to aid the scheduling scheme to work efficiently.

In addition, to reduce transmissions that are most likely to be unsuccessful, MAC can predict the packet transmission time based on the buffer size, average packet transmission times of different packet types, and their corresponding mean squared deviations. Then, MAC can compare the packet predicted transmission time with the packet network timeout period, the traffic delay bound, and the traffic delay jitter bound, so that it can notify the network layer whether the packet will timeout in advance.

B. Channel Condition Estimation and Packet Transmission Time Prediction at the MAC Layer

As described above, the MAC layer needs to estimate the channel condition and predict the packet transmission time.

In our scheme, we estimate the channel condition through monitoring the size changes of the buffer which stores MAC service data unit (MSDU) packets. The size is increased or decreased by one at one time. We denote the changes as a binary sequence. For example, during one specified period, the size change sequence may be as follows,

$$+1, +1, -1, +1, +1, +1, -1 .$$

Thus, if the initial buffer size is zero, then the last buffer size is

$$+1 + 1 - 1 + 1 + 1 + 1 - 1 = 3 .$$

We use three channel states to represent the channel conditions: busy, normal, and idle, defined as follows.

Definition 1. If there are more than C_1 consecutive $+1$ in the change sequence or there are more than C_2 MSDU packets in the buffer, we denote the channel state as busy.

Definition 2. If there are no packets in the buffer, we denote the channel state as idle.

Definition 3. The channel remains in the normal state if the channel is neither busy nor idle.

Once the channel state changes, the MAC layer will notify the network layer about the channel state immediately.

In order to predict the packet transmission time, we need the average packet transmission time and its corresponding mean squared deviation of each packet type. There are three packet types for MSDU: best-effort traffic packets, time-bounded traffic packets, and routing control packets (including route request, route reply and route error packets). For packet type t,

779

T_i and D_i are used to denote the average delay and delay jitter, respectively, which are computed as follows,

$$T_i = \frac{1}{N}\sum_{i=1}^{N} T_{ii} , \qquad (1)$$

$$D_i = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(T_{ii} - T_i)^2} , \qquad (2)$$

where N denotes the number of packets that the station has transmitted, T_{ii} represents the transmission time of the ith packet. In the following, we use b, v, and c as the subscript to represent the best-effort, time-bounded and routing control packet types. Actually, the packet transmission time includes waiting time, deferral time, back off time and actual transmission time.

Assuming that there are n packets remained in the MAC buffer, where $n = n_b + n_v + n_c$. As soon as a packet arrives from the network layer, the MAC layer computes the predicted transmission time of the packet through the following expression,

$$T_p = n_b \times (T_b + 2D_b) + n_v \times (T_v + 2D_v) + n_c \times (T_c + 2D_c). \quad (3)$$

C. Scheduling and Optimization Scheme

Once a packet arrives at the network layer, it will be put into a FIFO according to its traffic type. Then the station will schedule packet transmissions based on the different channel states estimated by the MAC layer. The scheduling algorithm is as follows.

When the channel condition is busy, the network layer will transmit the best-effort packets only after finishing the transmissions of all time-bounded packets. In normal state, if the time-bounded FIFO size is smaller than C_r, then the network layer will transmit the packets according to their insertion time into the two FIFOs. Otherwise, the station will transmit the packets in the time-bounded FIFO first. The station employs round robin to transmit the packets in FIFOs in the idle state.

In order to improve service differentiation and network performance, we adopt the following two optimizations in our scheme.

The first is the traffic properties sharing among the application, network, and MAC layers described above. At the network layer, when the station reads one packet from a FIFO, it first checks whether the packet has violated its delay and delay jitter bounds. If either bound is violated, it will be discarded. Otherwise, it will be transmitted. The MAC layer takes similar actions for different packets. That is, the network and MAC layers can both identify the different traffics of one packet through reading the TOS or TC field. At the application layer, once a packet arrives from the network layer, the station can determine whether to receive or discard it based on the delay and delay jitter bounds. Thus, the time delay and delay jitter can be guaranteed.

The second optimization is the interaction between the network and MAC layers. In IEEE 802.11, every station needs to acknowledge every MAC protocol data unit (MPDU) (e.g. RTS/CTS/DATA/ACK or DATA/ACK data transmission process). Whenever the MAC layer succeeds or fails to transmit a packet, it will notify the network layer immediately. This is similar to the acknowledgement process in the DSR model of [11]. In addition, in our scheme, MAC can notify the network layer about whether the packet has timed out or not in advance without actual transmission based on the packet transmission prediction value described above[1]. After getting the predicted value, the MAC layer will take the actions shown in Fig. 2.

```
If (traffic_type = best_effort)
  {
  if (Tp >= Trb - (current_time - packet_rcv_time))
    {
    discard the current packet;
    send timeout message to routing layer;
    }
  else
    {
    transmit the current packet;
    }
  }
else if (traffic_type = time_bounded)
  {
  if ((Tp >= Trv -- (current_time - packet_rcv_time)) or
               (Tp >= Td+Tj))
    {
    discard the current packet;
    send timeout message to routing layer;
    }
  else
    {
    transmit the current packet;
    }
  }
```

Fig. 2. Actions taken by the MAC layer after reading a packet from buffer

Through the above two optimizations, we can reduce the transmissions that are useless for the destination station. Thus, the contention is reduced, which is important for lightening the network load. In addition, this can reduce the memory requirements for the station because the optimizations eliminate storing useless packets in the FIFOs.

III. PERFORMANCE EVALUATION

A. Evaluation Model

We evaluate the scheduling and optimization scheme through simulations. We simulate 50 stations moving in a $600m \times 600m$ area according to the modified "random way-point" model [12] with the maximum and minimum moving speeds of 10m/s and 2m/s, respectively. We assume that every station has the capability to transmit and receive radio signals in a radius of 150m. Each source station randomly selects one of the 50 stations as its destination. The simulation period is 500 seconds. We use IEEE 802.11 MAC protocol (DATA/ACK mode) with 2Mbps channel capacity and DSR routing protocol based on [11].

---

[1] We have constructed (3) so that it is unlikely the transmission will be unsuccessful. In case it is underestimated, some packets will be discarded unnecessarily. This will not affect the overall performance.

We employ the following traffic generation models in the simulations. The packets of best-effort traffic are generated as a Poisson process with mean λ. The packet size is fixed. The ON-OFF model is used to simulate time-bounded traffic with the time parameters T_{on} and T_{off}. During the ON period, the packets are generated according to constant distribution with fixed packet size and its inter-arrival time is T_u. The packet transmission delay and delay jitter bounds are set to be T_d and T_j for time-bounded traffic.

The parameters used in the simulations are typical of the values used in the literature for IP voice and best-effort traffics. They are set to be $T_d = 0.25s$, $T_j = 0.025s$, $T_{on} = 1s$, $T_{off} = 1.35s$, $T_u = 0.1s$, $T_{rb} = 5s$, $T_{rv} = 0.25s$, $C_1 = 4$, $C_2 = 7$ and $C_r = 5$, respectively. The time-bounded packet size is fixed as 512 bits and the best-effort packet is 4096 bits. The delay and delay jitter bounds of best-effort traffic are both set to ∞.

We validate our scheme under different network loads. In the simulations, there are three types of sources: best-effort, time-bounded, and combined sources (i.e., transmitting best-effort and time-bounded packets simultaneously). For comparison, we also simulate the above-mentioned IEEE 802.11 MAC and DSR protocols under the same conditions, but not implementing the proposed scheduling and optimization scheme. The performance evaluation metrics are as follows.

- *Packet transmission average delay*: the average delay a packet takes from its generation at the source to arriving at the destination.
- *Packet transmission delay jitter:* the mean squared deviation of packet transmission average delay.
- *Packet delivery ratio: the ratio of the correctly received packets at the destination station to the total generated packets by the source stations.*

B. Simulation results

In the following, for ease of description, "SO Scheme" and "Non-SO Scheme" are used to denote the schemes employing and not employing the proposed scheduling and optimizations, respectively.

Fig. 3 and Fig. 4 show the average delay and delay jitter when there are 5 best-effort sources, 5 time-bounded sources and 5 combined sources in the network. It is clear that the time-bounded delay and delay jitter of "SO Scheme" are much lower than that of the others. This is due to the higher priority transmissions of time-bounded packets and the optimizations, which include discarding those packets whose transmission times exceed the sum of the predefined time delay and delay jitter bounds.

It should also be noted that the delay jitters of best-effort packets of both the "SO Scheme" and "Non-SO Scheme," and the delay jitters of the time-bounded packets of the "Non-SO Scheme" are bigger than their corresponding delays. The reason is that although most of the packets arrive at the destinations in a

short time, some of them will spend much longer time to arrive due to channel contentions and frequent topology changes. Fig. 5 shows such an example, in which the x-axis represents the simulation time from 300s to 500s. Clearly, the impact of several large packet transmission delays can be smoothed when computing the transmission average delay. But their impact on the delay jitter is huge.

At the same time, from Fig.3, there is a difference between the average delays of best-effort and time-bounded traffics of the "Non-SO Scheme." The reason is two-fold. One is the different packet lengths induced. From [6], different packet lengths can have different packet transmission delays due to the different packet collision probabilities and different wireless transmission times. The other is the different characteristics of traffic packet generation. Since the time-bounded traffic packet generation follows the constant distribution during the ON state, the routes are frequently and continuously used and updated, leading to fresher routes. Whereas, for the best-effort packet transmission, the route may be obsolete when the next packet arrives because of the longer inter-arrival time of best-effort packets. Thus, the station has to initiate the route request process to find another feasible route. The above reasons lead to the bigger average delay of best-effort packets.
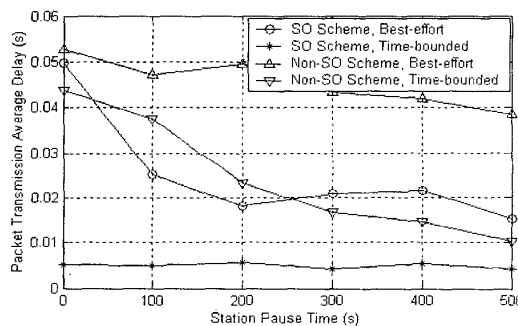

Fig. 3. Packet transmission average delay vs. station pause time

The packet delivery ratios of both the "SO Scheme" and "Non-SO Scheme" are shown in Fig. 6. In the simulations, there are 10 best-effort sources, 10 time-bounded sources, and 10 combined sources. Obviously, the packet delivery ratio of time-bounded traffic of the "SO Scheme" is relatively stable and high under different station pause times. Furthermore, either the time-bounded or the best-effort packet delivery ratio of "SO Scheme" is higher than that of the "Non-SO Scheme" when the station pause time exceeds 60s. This is due to the reduced packet collisions and network load resulting from the proposed optimizations, which avoid transmitting packets that are useless for the destinations and are most likely to be unsuccessful when transmitted. Similarly, the difference of the delivery ratios of best-effort and time-bounded traffics of the "Non-SO Scheme" is caused by the same reasons described above.

It is known that for IP voice traffic, when the delay bound is about 0.25s, the delay jitter bound is about 0.025s, and the packet loss ratio is between 5%~10%, the voice quality is acceptable. From the above simulation results, it is evident that our scheme can satisfy these requirements.
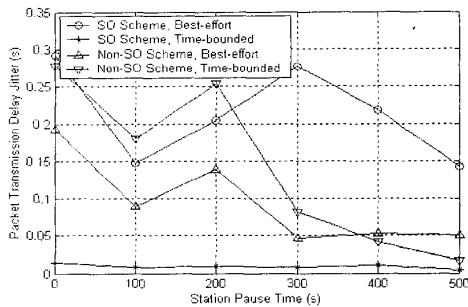
781

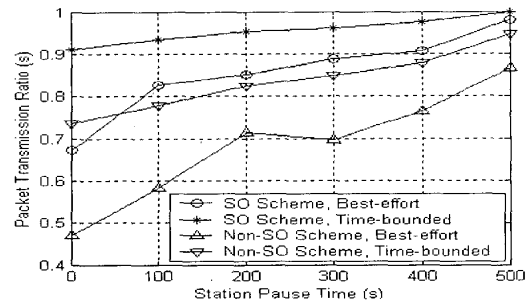Fig. 4. Packet transmission delay jitter vs. station pause time
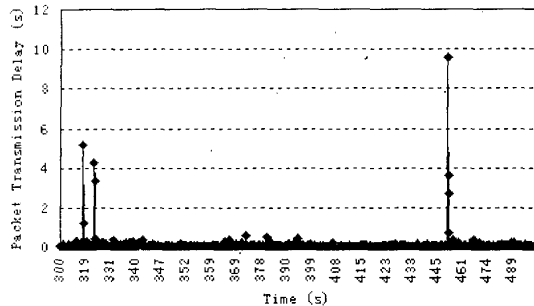

Fig. 5. Packet transmission delay

IV. CONCLUSIONS

In this paper, we employ cross layer design to develop a scheduling and optimization scheme for service differentiation in MANET. Through monitoring the size changes of the buffer which stores the packets arriving from the network layer, the MAC layer determines the channel state and behaves accordingly. Then, this information is conveyed to the network layer to aid packet transmission scheduling. The optimizations are based on traffic property sharing and packet timeout period interaction between the network and MAC layers to reduce the packet collisions and improve network performance. By employing the above optimizations, the transmitted packets that are useless for the destinations and are most likely going to be unsuccessful are reduced. This decreases packet collisions and improves the wireless channel bandwidth efficiency.

To evaluate our proposed scheme, we simulate different network conditions. The simulation results .show that our scheme can provide service differentiation for best-effort and time-bounded traffics. IP voice traffic is very sensitive to the time delay and delay jitter. We show that our scheme can satisfy its requirements.

Fig. 6 Packet delivery ratio vs. station pause time

REFERENCES

[1] J. Broch, D. B. Johnson, D. A. Maltz. "The Dynamic Source Routing protocol for mobile ad hoc networks." Internet Draft, draft-ietf-manet-dsr-*.txt.

[2] W. H. Yuen, H. Lee, T. D. Andersen, "A simple and effective cross layer networking system for mobile ad hoc networks," IEEE PIMRC, Lisboa, Portugal, September 2002.

[3] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad-hoc networks," IEEE InfoCom, New York, NY, USA, June 2002.

[4] B. G. Chun, D. R. Jaisinghani, M. Baker, "Evaluation of scheduling algorithms in mobile ad hoc networks," ACM MobiHoc, Lausanne, Switzerland, June 2002.

[5] X. Pallot, L. E. Miller, "Implementing message priority policies over an 802.11 based mobile ad hoc network," IEEE MILCOM, Washington, D.C., USA, October 2001.

[6] I. Aad, C. Castelluccia, "Differentiation mechanisms for IEEE 802.11," IEEE InfoCom, Anchorage, Alasaka, USA, April 2001.

[7] X. Yang, N. H. Yaidya, "Priority scheduling in wireless ad hoc networks," ACM MobiHoc, Lausanne, Switzerland, June 2002.

[8] H. Luo, S. Lu, V. Bharghavan, "A new model for packet scheduling in multihop wireless networks," ACM MobiCom, Boston, Massachusetts, USA, August 2000.

[9] S. Lu, V. Bharghavan, R. Srikant, "Fair scheduling in wireless packet networks," IEEE/ACM Transactions on Networking , Volume 7, Issue 4 August 1999

[10] A. Dugar, N. Vaidya, P. Bahl, "Priority and fair scheduling in a wireless LAN," Technical Report TR01-006, Computer Science Department, Texas A&M University, April 7, 2001.

[11] X. Pallot, N. Roux, J. S. Pegon, "README File for NIST DSR Model," File and Model found at "http://w3.antd.nist.gov/wctg/DSRreadme.pdf," December 2000.

[12] J. Yoon, M. Liu, B. Noble, "Random waypoint considered harmful," IEEE InfoCom, San Francisco, California, USA, March 2003

782