Supporting Service Differentiation for Real-Time and Best-Effort Traffic in Stateless Wireless Ad-Hoc Networks (SWAN)

G. S. Ahn, A. T. Campbell, A. Veres, and L. H. Sun


One of excellent papers for service differentiation provisioning over mobile wireless ad-hoc networks
Outline

- Motivation
- Related Work & Main Idea
- Distributed Control Algorithm
- Experimental Results
- Conclusion and limitations
Motivation

Provide service differentiation for real-time traffic and best effort traffic in mobile ad-hoc wireless networks in a simple, scalable, and robust manner.
Basic Assumption

Most of the wireless network capacity will be utilized by best effort traffic.

Best effort traffic can be used as a “Buffer Zone” to absorb real-time traffic variation due to mobility or traffic bursty, i.e., reduce best effort traffic rate when real-time traffic is bursty or heavy.
Related Work & Main Idea

- AIMD (additive increase multiplicative decrease) for congestion avoidance

- Endpoint admission control for wireline networks
  - Stateless, scalability and providing "soft" quality of service
  - Periodically send probe packets to measure available bandwidth for traffic flows
Main Result

- A simple, scalable, and robust architecture to deliver service differentiation in mobile wireless ad-hoc networks
  - **Simple**: only use classic MAC protocol
  - **Scalable**: stateless, local control mechanism
  - **Robust**: changes in network topology, real-time traffic load, and link failures do not affect the operation of proposed control system
Distributed Control Algorithms
(Core Part in SWAN)
Two Components in Distributed Control Algorithms

- **Classifier**
  - Differentiate real-time and best effort traffic and force traffic shaper only process best effort packets

- **Shaper**
  - Simple leaky bucket traffic control to delay best effort traffic packets in conformance with the available rate
Rate controller

- Only process best effort traffic packets
- Each node in mobile ad hoc network independently performs rate control
- Use local information
  - Packet delay measured at MAC layer
  - Delay = Ack receiving time - Packet arrival time
- Rate controller determines the allowed transmission rate $s$ for best effort traffic using AIMD rate control algorithm
  - Every $T$ seconds
    - Increase $s$ by $c$ Kbps if packet delay < predefined delay bound $d$ sec to efficiently utilize resource
    - Decrease $s$ by $r$ percent if packet delay > $d$ to guarantee real-time delay bound requirement
    - Decrease $s$ to $a^*(1+g/100)$ if $s \gg$ actual best effort traffic rate $a$ to control bursty of best effort traffic
Procedure update_shaping_rate ( )
/* called every T second period */
Begin
if (n > 0) /* one or more packets have delays
greater than the threshold delay d sec */
s ← s * (1 - r / 100) /* multiplicative decrease by r% */

else
s ← s + c /* additive increase by c Kbps */

if ( (s - a) > a * g / 100) /* difference between actual rate and shaping
rate is greater than g% of actual rate */
s ← a * (1 + g / 100) /* adjust shaping rate to match actual rate */
end
Admission Controller

- Only perform at traffic source node
- Threshold rate
  - Rate that would trigger excessive delays
- Admission control rate
  - Rate for real-time traffic
- Conservative admission control rate
  - e.g., IEEE 802.11b 11 Mbps
  - Threshold rate = 3.5 Mbps
  - Admission control rate = 2 Mbps
- Exploit the broadcast nature of wireless communication to estimate local real-time traffic rate
- Available bandwidth for real-time traffic
  - admission control rate - current rate of local real time traffic
Admitting A New Real-Time Session

- Source sends a probing request packet with a “bottleneck bandwidth” field to estimate end-to-end bandwidth availability.

- Intermediate nodes update the field if current value in packet header > bandwidth availability at the node.

- Destination node sends a probing response packet back to source node with the bottleneck field copied from the probing request packet.

- Source node accepts new real-time session if end-to-end available bandwidth > required bandwidth by new real-time session.
Challenges for Proposed Admission Control Algorithm

- Two scenarios can cause excessive delay in real-time traffic by exceeding the threshold rate
  - Mobility
  - False Admission
    - Multiple source nodes simultaneously sending probing request packets and share some common nodes
    - Available bandwidth is shared multiple source nodes, but treated as for every source node
- Need regulation algorithms to overcome
Dynamic Regulation Algorithms

ECN based regulation
- Each node periodically measures utilization of real time traffic
  - If local aggregated real-time traffic rate > conservative admission control rate
  - Marks ECN bit in IP header of the real time packets
- Destination node monitors the ECN bits and informs the source via regulate message

Existing problem
- If all real-time packets are marked with ECN, all real-time sessions would be forced to reestablish their service simultaneously

Two proposed approaches
- Source based regulation
- Network based regulation
Source Based Regulation

- Source node can differentiate the regulation associated with false admission or mobility by using state information about newly admitted vs ongoing flows
  - Take action immediately for new admitted real-time session
  - Otherwise wait a random amount of time before starting the reestablishment

- Advantage
  - Purely source based

- Disadvantage
  - Sources that regulate earlier may find the path to be overloaded and be forced to drop their sessions
Network Based Regulation

- Rather than marking all real-time packets with ECN bit, the intermediate node selects a congestion set and mark real-time packets associated with the set.

- Falsely admitted flows are distinguished using an additional bit in the ToS field to check whether the real-time session is old or new.

- Disadvantage
  - Intelligence needed at intermediate nodes to select a random congestion set.
Comparison between Source Based and Network Based Regulation

At time 20 sec, rerouting occurs due to mobility and new real-time sessions arrive.

Single Broadcast region
Channel Bandwidth 11 Mbps
Threshold rate 3 Mbps
Admission control rate 2 Mbps

Network-based regulation performs better at cost of 1-bit in packet header to mark new/old session and intelligence at intermediate nodes to select a random congestion set.

At time 20 sec, rerouting occurs due to mobility and new real-time sessions arrive.
Simulation Setting for Single Shared Channel (One Broadcast Region)

- Transmission range 250 meter
- Radio channel bandwidth 11 Mbps
- Simulated area 150 m by 150 m
- TCP flows: greedy ftp traffic with 512 bytes packet size
- UDP flows
  - Voice traffic: 32 Kbps with 80 Bytes packet size
  - Video traffic: 200 Kbps with 512 bytes packet size
- Real-time traffic: 4 voice flows and 4 video flows
- Best effort traffic: up to 32 TCP flows
Evaluation and comparison of performance of SWAN, DCF and CWmin
Performance of a single shared channel

Impacts of AIMD parameter (c, r)

Average delay of real-time traffic versus increment rate.
Total throughput of best-effort TCP traffic versus increment rate.
Evaluation and comparison of performance of SWAN, DCF and CWmin

Performance of a single shared channel

Impacts of AIMD parameter \((c, r)\)

When \(r = 50\), delay is reduced 60 percent for 8 TCP flows, 75 percent for 32 TCP flows at cost of no more than 5 percent TCP throughput loss.

Average delay of real-time traffic versus decrement rate. Total throughput of best-effort TCP traffic versus decrement rate.
Comparison of DCF, CWmin and SWAN

Real-time traffic: 4 voice flows and 4 video flows

SWAN: up to 88 percent delay reduction at cost of at most 2 percent TCP throughput loss

1. Average delay of real-time traffic versus number of TCP sources.
2. Total throughput of best effort TCP traffic versus number of TCP sources.
Comparison of DCF, CWmin, SWAN-RC and SWAN

Single broadcast region, 16 TCP flows

Fig. 13. Average delay of real-time traffic versus number of video sources.
Simulation Setting for Multihop Scenarios

- Transmission range 250 meter
- Radio channel bandwidth 11 Mbps
- Simulated area 1500 m by 300 m
- 50 nodes
- AODV routing protocol
- TCP flows: greedy ftp traffic with 512 bytes packet size
- UDP flows
  - Voice traffic: 32 Kbps with 80 Bytes packet size
  - Video traffic: 200 Kbps with 512 bytes packet size
- Real-time traffic: 4 voice flows and 4 video flows
- Best effort traffic: up to 12 TCP flows
Performance of Multihop Scenarios with Mobility

38-77 percent delay reduction at cost of 15-20 percent throughput loss

Average delay of real-time traffic versus number of TCP flows. 5. Average “goodput” of TCP best-effort traffic versus number of...
Performance of Multihop Scenarios with Mobility

Random waypoint mobility model

70-75 percent delay is reduced at cost 15-25 percent throughput loss

Average delay of the real-time traffic versus mobility.

Average goodput of the best-effort TCP traffic versus mobility.
Conclusion and limitation

- An simple and efficient architecture (SWAN) for mobile wireless ad-hoc network to support real-time and best effort traffic flows

Limitations
- how do determine parameters?
  - r (multiplicative decrease)
  - c (additive increase)
  - g (difference between the actual rate and the shaping rate)
- admission control rate