Increasing Broadcast Reliability for Vehicular Ad Hoc Networks

Nathan Balon
Advisor: Dr. Jinhua Guo

University of Michigan - Dearborn

Summary

- General Information on VANETs
- Background on 802.11
- Background on 802.11e for QoS
- The issues involved with transmitting a broadcast message in a VANET
- A modified CW algorithm to increase the reception rate of broadcast packets
- Conclusion

Vehicular Ad Hoc Networks

- The goal of a VANET is to increase the overall safety of the transportation infrastructure.
- Automobile accidents are one of the leading causes of death in the US. To illustrate the severity of the problem of automobile accidents, there were 6,279,000 traffic accidents that accounted for 41,611 deaths in the United States in 1999.
- In 1999, the FCC allocated 75 MHz of bandwidth in the 5.9 GHz band to create a nationwide VANET.

Dedicated Short-Range Communication

- Dedicated Short-Range Communication (DSRC) refers to the spectrum allocated by the FCC for vehicular networks.
- The DSRC program's goal is to create an interoperable standard for use throughout North America (US, Mexico, and Canada).
- DSRC will allow drivers to receive up-to-date information regrading their driving environment and as a result reduce accidents.

Requirements for a VANET

- A VANET must support high data rates.
- The latency must be low, 100 *ms* or less.
- The communication range of VANET is relatively small; it ranges from 100 *m* to 1,000 *m*.
- DSRC is intended to support both public safety and licensed private operations.
- No fee should be required to access the the public safety service of the VANET.

IEEE 802.11p

- The IEEE standardization process is under way for the PHY and MAC layer of DSRC, and the new standard is known as 802.11p.
- The PHY layer of 802.11p is similar to 802.11a.
- As with 802.11a, the modulation scheme used for the PHY of 802.11p is Orthogonal Frequency Division Multiplexing (OFDM).

DSRC Channels

5.850 GHz		CH. 175			CH. 181		5.925 GHz
Reserved	CH. 172 Service	CH. 174 Service	CH. 176 Service	CH. 178 Control	CH. 180 Service	CH. 182 Service	CH. 184 Service
5 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz	10 MHz

Control Channel

- Channel 178 of DSRC is reserved for the control channel.
- The control channel is the most important channel of DSRC and efficient use of the channel is critical.
- The control channel is used to exchange service announcements, safety messages, and a vehicle's state.
- Communication on the control channel must last less than 200 *us*.
- Every 100 *ms*, each vehicle broadcasts its state on the control channel (e.g., direction, speed, location, etc.)

DSRC Data Rates

- The data rates possible for a 10 Mhz channels are:
  - 6, 9, 12, 18, 24, and 27 Mbps
- It is also possible to combine two 10 MHz channels to achieve a data rates up to 54 Mbps.

Communication Devices

- Each vehicle has an On Board Unit (OBU).
- The OBU consists of:
  - Transceiver
  - Omni-directional antenna
  - Processor
  - GPS unit
  - Digital Maps
  - Sensors

Communication Devices

- Road-Side Units (RSU) also exist within a VANET.
  - The RSU are strategically placed along the side of the road to provide services to vehicles.
  - The RSU's components are similar to those contained in an OBU.
  - A RSU can communicate with other RSUs through a wired infrastructure.
  - Each RSU will require a license to operate that unit at a specific location and a specific frequency.

VANET Applications

- Traffic Signal Violation Warning
- Curve Speed Warning
- Emergency Electronic Brake Lights
- Cooperative Collision Warning
- Left Turn Assistance
- Pre-crash Warning
- Stop Sign Movement Assistance

802.11

- The MAC layer for DSRC is based on 802.11.
- The 802.11 standard defines two MAC protocols.
  - Distributed Coordination Function (DCF)
  - Point Coordination Function (PCF)
- The DCF is an asynchronous contention based access protocol that is used with DSRC.
- In a contention based protocol, all nodes that have data to send contend for access to the channel.
- The PCF has contention free period during which the AP coordinates the transmission of the stations.

802.11

- 802.11 family of protocols uses Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA).
- The DCF achieves collision avoidance with a random back-off procedure.
- The IEEE 802.11 standard uses the concept of slot-time. For 802.11a time-slots are 9 *us*.
- When a node begins a transmission, it randomly selects the number of time-slots it must wait before transmitting, which is known as the back-off window (contention window).
- Since there is no collision detection, the wireless transmission is made reliable with the introduction of an explicit acknowledgement.

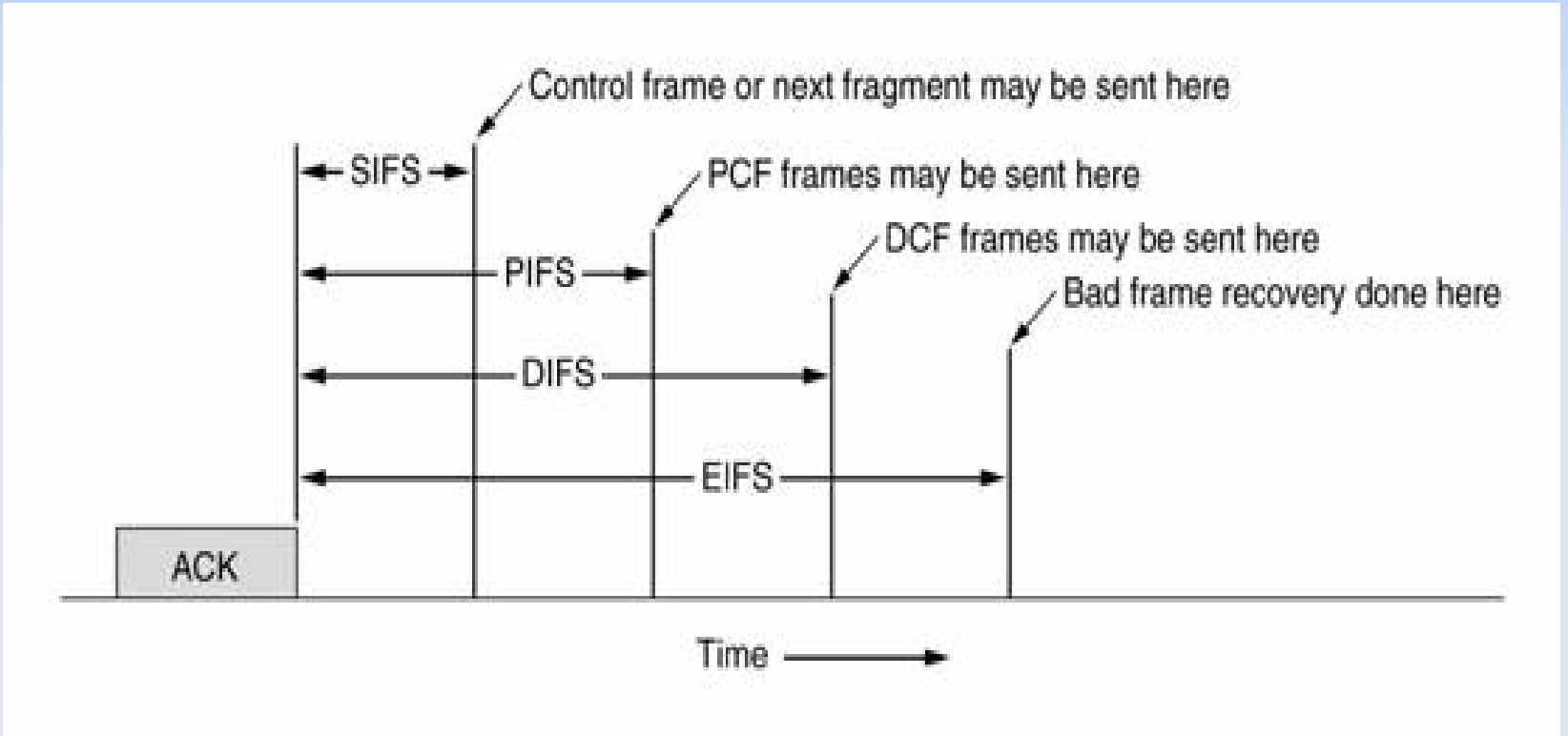
802.11: Inter-frame Spacing

- 802.11 uses a number of different inter-frame spaces.
- The Short Inter-Frame Space (SIFS) is the shortest inter-frame space used for 802.11. All of the other inter-frame spaces are defined in relation to SIFS. For instance, for 802.11a the value of SIFS is 16 μs.
- The Distributed Inter-Frame Spacing (DIFS) is used before the transmission of both broadcast and unicast frames. For 802.11a the length of DIFS = 2 x Slot-time + SIFS = 34 μs.

802.11: Inter-frame Spacing

- When a node wishes to transmit a frame, it must wait for the DIFS timer to expire before attempting to access the channel.
- During this time, the wireless medium must remain free of transmissions.
- If a transmission is overhead while a node is waiting for the DIFS expire, the node defers its attempted access to the medium until the medium becomes free and the DIFS timer is restarted.

Inter-Frame Spacing



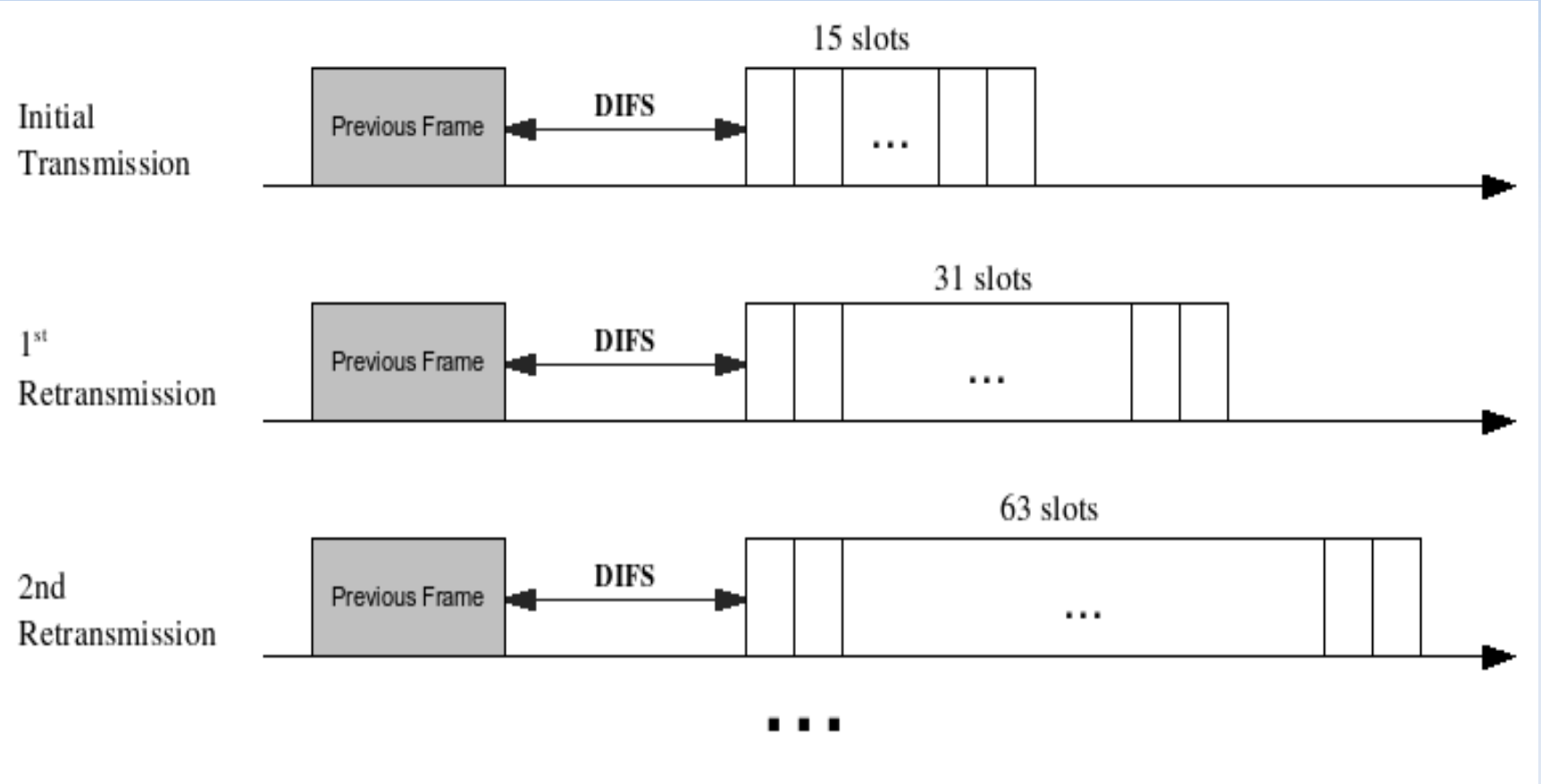
Back-off Process

- After the completion of the DIFS a node begins a back-off procedure.
- When a packet is passed down to the MAC, it randomly selects the number of time slots it must wait before transmitting the packet; this is known as the contention window.
- One clock tick of the back-off timer expires, when the medium remains free from transmission for one time-slot.
- The back-off timer is paused if the medium becomes busy.
- A node transmits when its back-off timer reaches zero.

Back-off Process

- The initial value of the CW is chosen randomly from the interval [0, CW_{min}].
- If the transmission fails the CW is exponentially increased until CW_{max} is reached.
- Some typical CW Sizes are: 7, 15, 31, 63, 127, 255, 511, 1023
- If the transmission of a frame does not succeed after a predefined numbers of attempts the frame is discarded.

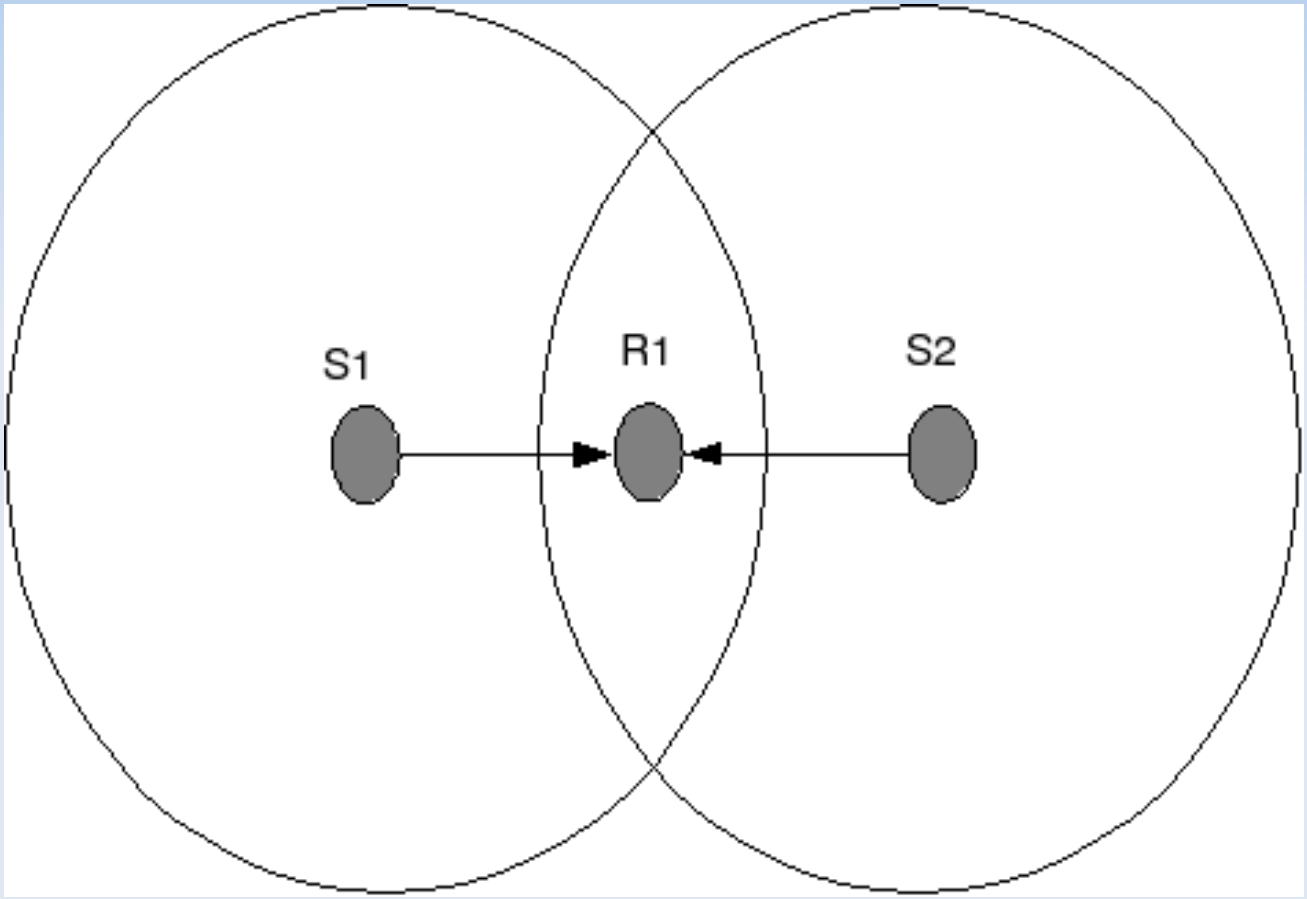
802.11 Retransmission Attempts



Hidden Terminal Problem

- One of the main problems affecting the reliability of DCF is the hidden terminal problem.
- The hidden terminal problem occurs when there are two nodes that are outside of the transmission range of each other, but each node wishes to transmit to a node that is shared between them.
- The problem is the nodes cannot detect each others transmissions, so the medium appears free to both nodes.
- If Both nodes transmit at the same time the frame will fail to be delivered, and a collision will occur at the receiver.

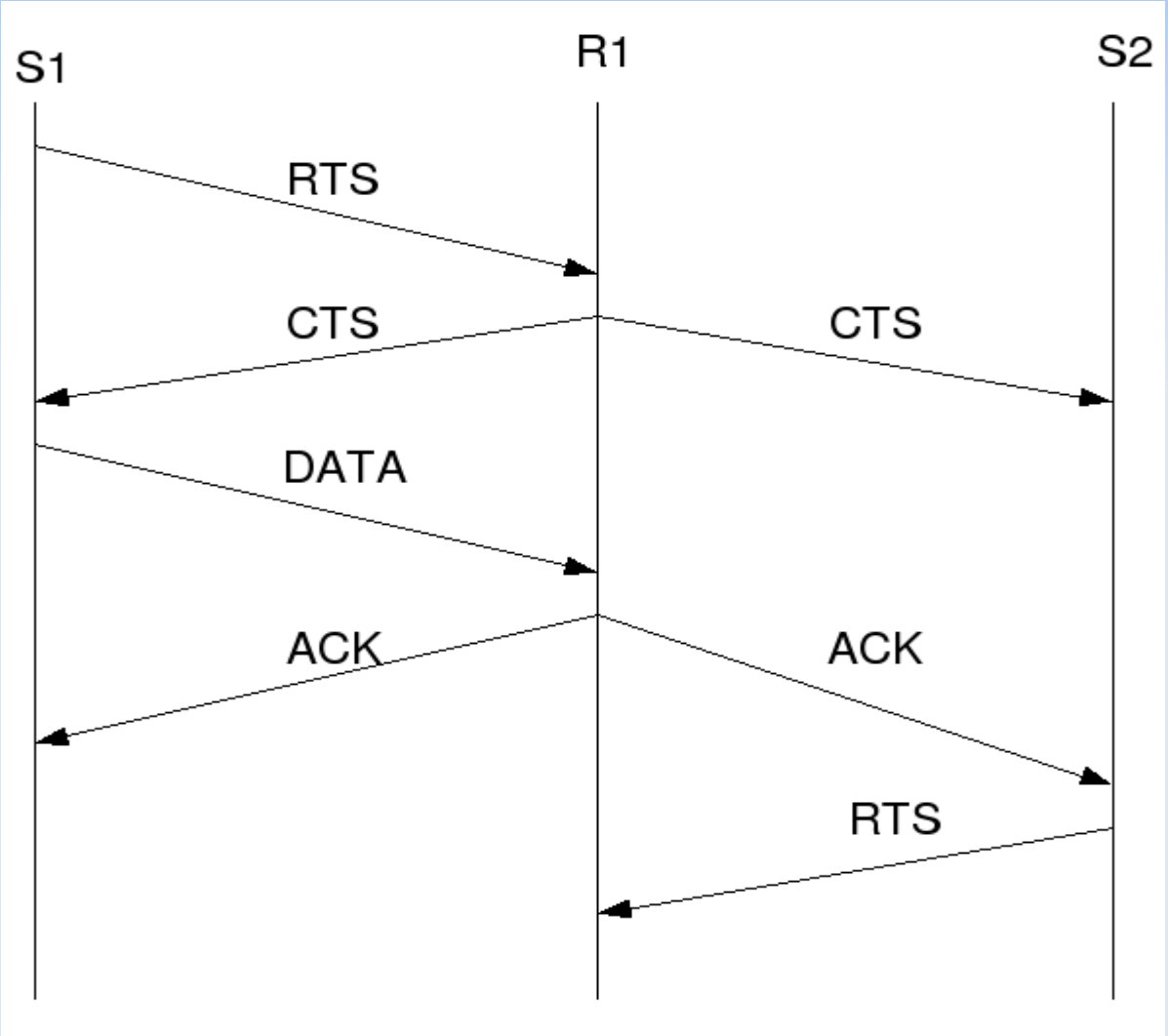
Hidden Terminal Problem



RTS/CTS

- To overcome the hidden terminal problem an optional Request to Send/Clear to Send (RTS/CTS) exchange was introduced into 802.11.
- When a node has data to send, once it is able to gain access to the medium, the node will first transmit a RTS to the intended receiver.
- When the intended target receives the RTS, after a SIFS has expired, the node will respond with a CTS.
- All nodes that overhear these control messages set their network allocation vector (NAV) for the amount of time it will take to complete the communication.
- These other nodes will then defer from accessing the wireless medium until their NAV expire, and the transmission between sender and receiver is complete.

RTS/CTS



802.11e QoS

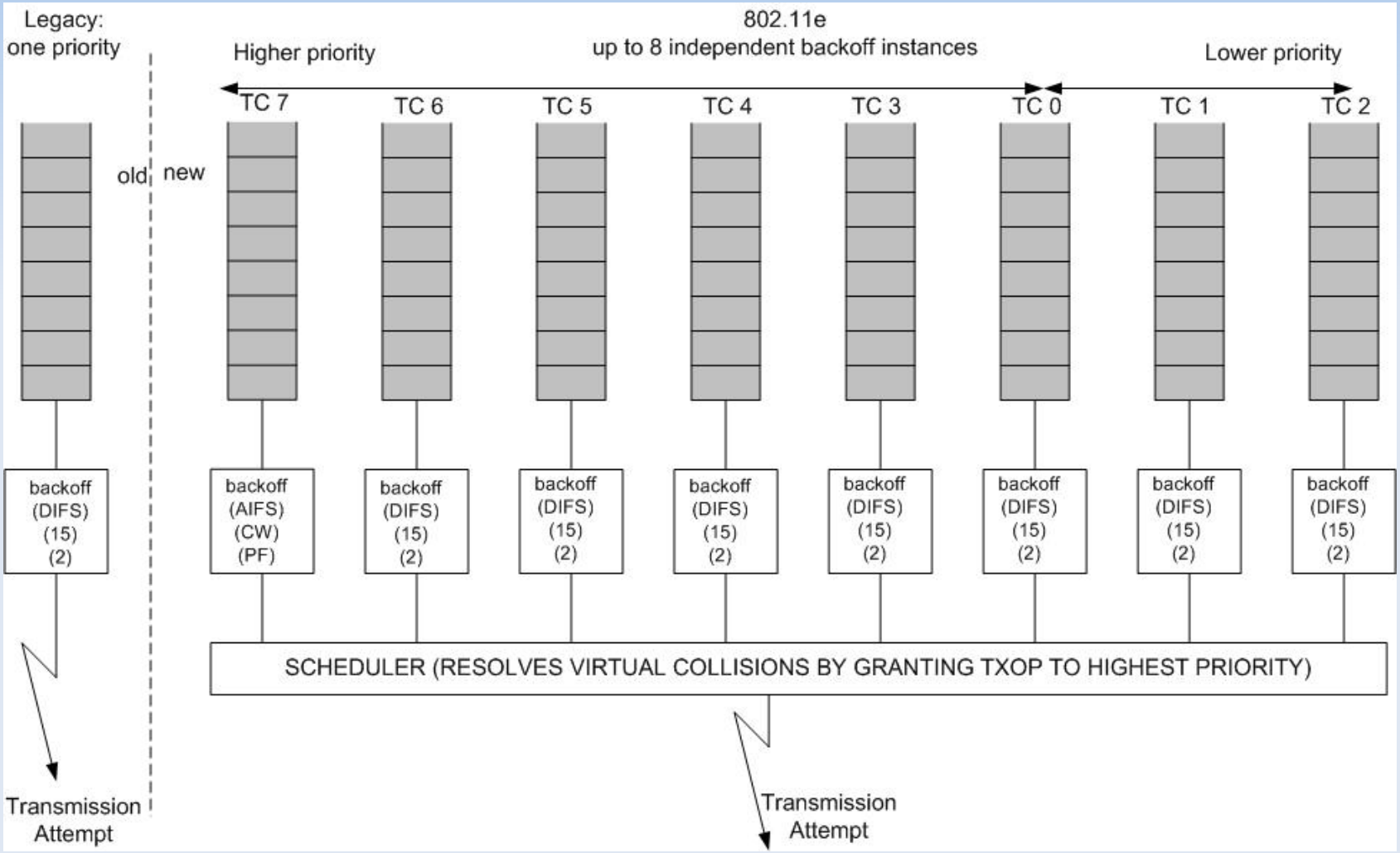
- A requirement of DSRC is that safety messages must have priority access over non-safety messages.
- In order to timely deliver high priority messages, such as those used by collision warning applications, DSRC adopts the Enhanced Distributed Channel Access (EDCA) of 802.11e.
- EDCA provides differential access to the wireless channel by using eight priority classes, which are referred to as Access Categories (AC).
- Each AC has a different set of access parameters.

Access Categories

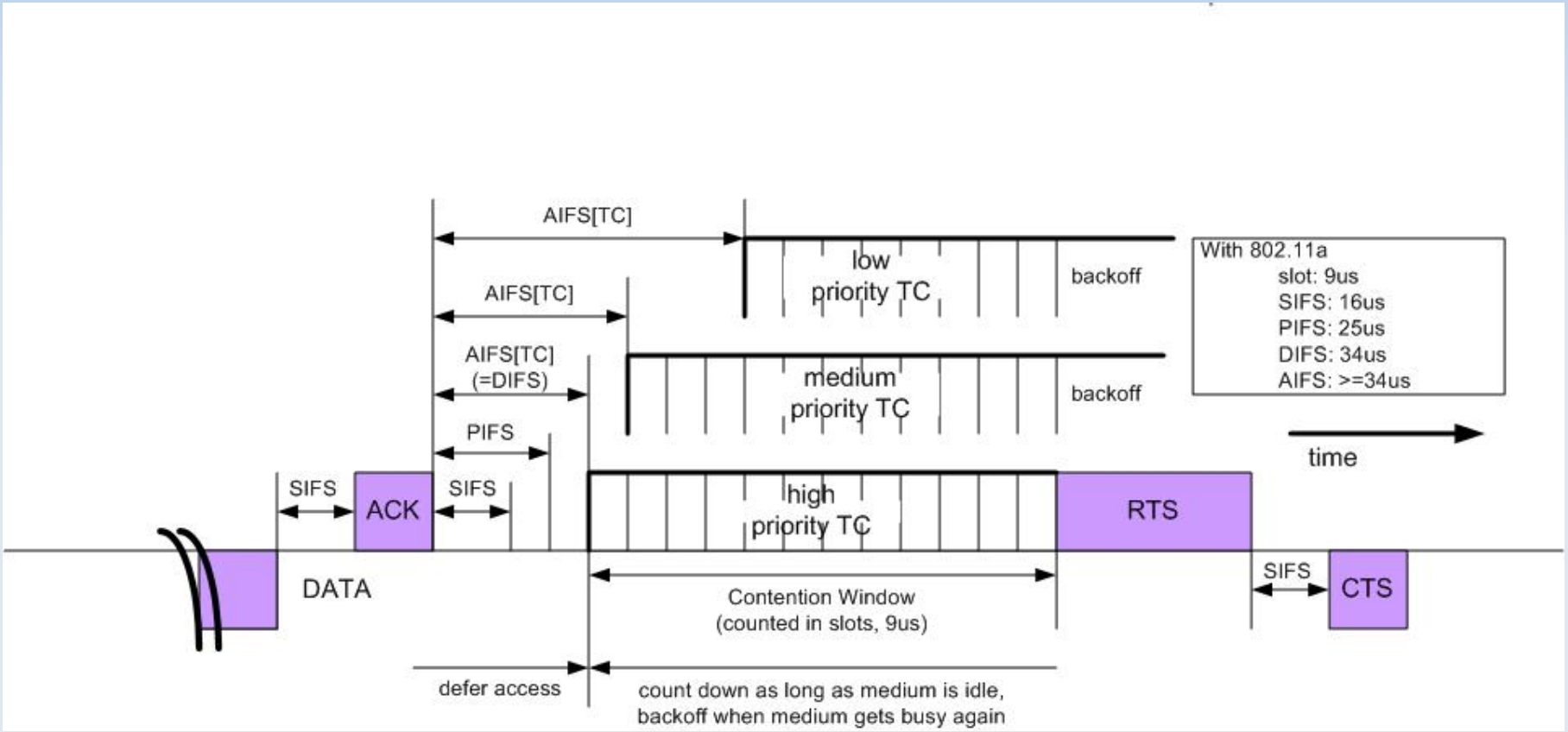
Table 3.4: Access Categories

AC	$AIFS[AC]$	$CW_{min}[AC]$	$CW_{max}[AC]$
0	2	7	15
1	2	15	63
2	3	15	511
3	4	31	1023

802.11e Queues



802.11e Backoff



Broadcast Messages

- Broadcast messages will play a larger role than unicast messages in a vehicular environment.
- Some uses of broadcasting are, as explained earlier, to send emergency warning messages and to periodically broadcast a vehicle's state.
- When a frame is broadcast no ACK or RTS/CTS control frames are used.

Problems with Broadcasts

- No retransmission is possible for failed broadcast transmissions because of the lack of explicit acknowledgement for broadcast frames.
  - A failed unicast transmission is detected when an acknowledgment (ACK) is not received at the transmitter.
  - If acknowledgments were used for broadcasts, a problem known as the "ACK explosion problem" would exist. Each receiving node would at almost the same instance send an ACK back to the transmitter.
  - The "ACK explosion problem" results in a large number of collisions occurring.

Problems with Broadcasts

- The hidden terminal problem exists because the RTS/CTS exchange is not used.
  - The IEEE 802.11 protocols use an optional RTS/CTS handshake followed by an acknowledgment to guarantee the delivery of a unicast packet.
  - Broadcast messages, on the other hand, cannot use the RTS/CTS exchange because it would flood the network with traffic.
  - All nodes that received the RTS would at almost the same instant respond causing a storm around the node that transmitted the RTS.
  - Because the RTS/CTS is not used channel reservation is not possible.

Problems with Broadcasts

- The contention window size fails to change because there is no MAC-level recovery on broadcast frames.
  - In order to control congestion, the contention window size (CW) is exponentially increased each time a failed unicast transmission is detected.
  - Because there is no detection of failed broadcast transmissions, the size of the CW fails to change for broadcast traffic as it does for unicast traffic.
  - Nodes will always transmit with CW_{min} for the backoff window.

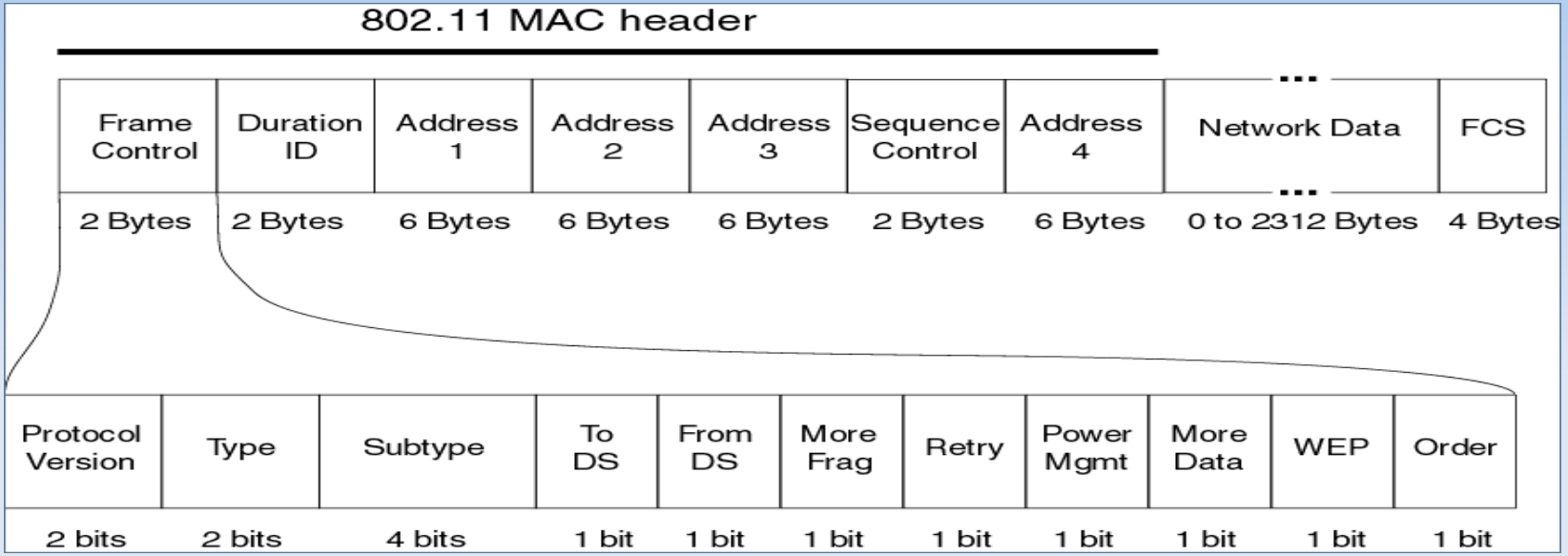
Modified Broadcast Protocol

- We rely on the observation that a node in a VANET is able to detect collisions and congestion by simply analyzing the sequence numbers of packets it has recently received.
- In a VANET, each node will broadcast its status to its neighbors approximately 10 times every second.
- While a node does not know if the packets it sent are correctly delivered, it knows the exact percentage of packets sent to it from its neighbors that are successfully received.
- Based on this feedback, a node is able to dynamically adjust the parameters it uses, such as contention window size, transmission rate, and transmission power, to improve the delivery rate of broadcast messages.

Modified Broadcast Protocol

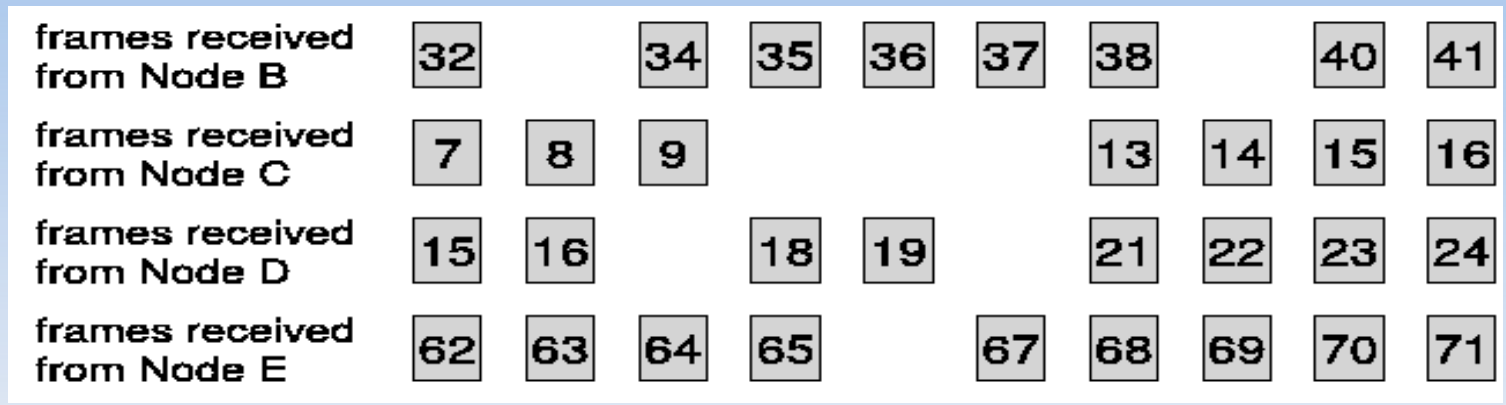
- The probability of collisions can be reduced and probability of reception can be improved if the size of the CW used to send broadcast messages is able to adapt based on the network conditions.
- If a node is able to observe its local conditions, it will be able to modify the MAC level parameters to improve the probability that a frame is successfully received.
- While it is not possible to detect the collisions of frames, it is possible to record the successful delivery of frames.

802.11 MAC Header



- The sequence control field is 2 Bytes.
  - A 12-bit sequence number is contained within the sequence control field.
  - It result in modulo-4096 counter that is incremented by 1 for each packet passed to the MAC.

Frames Received at Node *A*



- Node *A* records that it has overheard the frames coming from node *B* with the sequence numbers: 32, 34, 35, 36, 37, 38, 40, and 41.
- Based on the observed sequence numbers, node *A* could conclude that frames 33 and 39 coming from node *B* were corrupted or lost.
- Node A can also conclude that frames 10, 11, and 12 were not correctly received from Node *C*.

Data Maintained by Each Node

- Each node then records the overheard sequence numbers coming from specific nodes.
- A dynamic hash table is used so that an entry is updated in near constant time (i.e., $O(1)$). The MAC address is used as the key to the hash function.
- Each table entry also has a time-stamp associated with it. In order to prevent old data from affecting the calculation of the local network conditions, old entries are periodically removed.

MAC Address	Sequence Number	Average Reception Rate	Timestamp

Weighted Moving Average

- A weighted reception rate is used to determine the percentage of packets that are successfully received from a specific node.
- The weighted reception rate is used to put more emphasis on recent events.
- Each time a frame is successfully received, the weighted reception rate is recalculated.
- The variable α is used to put more or less weight on the current network condition.
  - $\text{EstRecpRate} = \alpha * \text{EstRecpRate} + (1 - \alpha) * \text{SampRecpRate}$

Local Reception Rate

- The nodes also maintain a timer. When the timer expires, the local reception rate is determined and the CW is adjusted.
- The local reception rate is the average of the estimated reception rates.
  - $\text{LocalReceptionRate} = \sum \text{EstRecptionRate} / \text{Nodes}$

Contention Window Adjustment

- After the local reception rate is calculated, it is compared against the previously stored local reception.
- The node also maintains a threshold value.
- If the new local reception rate decreases by a value greater than the threshold, the CW is increased.
- On the other hand, if new local reception rate increases by a value greater than the threshold, the CW is decreased.
- As a result, the CW adapts to the condition of the network.
- The node then uses the new CW for all broadcast transmissions until the next time the local reception rate is calculated.

CW Adjustment

IF (average – previous average >= threshold)

Increase the window

ELSE IF (-(average – previous average) >= threshold)

Decrease the window

ELSE

Maintain the current window

Network Simulations

- The NS-2 network simulator was used to simulate the modified version of the 802.11 protocol.
- NS-2 is a discrete event simulator that is extensively used for network research.
- The simulator was developed at University of California Berkeley and it is written in both C++ and TCL.

NS-2 Simulations

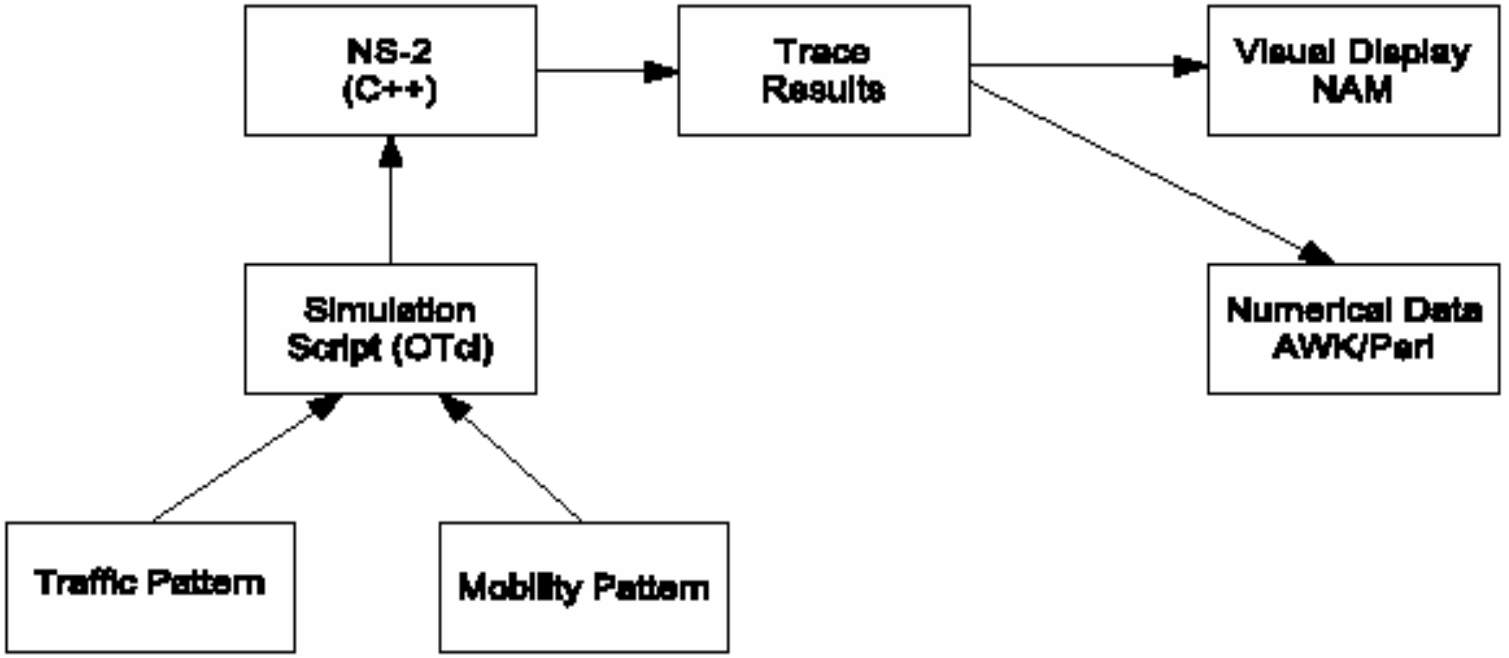


Figure 4.3: Steps Involved in an NS-2 Simulation

Mobility Model

- The mobility model used in the simulations is Freeway Mobility Model, and the USC Mobility Generator was used to create the mobility in the simulations.
  - Each node in the simulation is restricted to only travel within its lane.
  - The velocity of each node is temporally restricted based on the nodes previous velocity.
  - A safety distance is maintained so that a node cannot exceed the velocity of the node in front of it if they are within the safety distance.
- A velocity range is specified for the nodes. For the simulation each vehicle's velocity ranges from 17m/s to 25m/s.
- The acceleration of the vehicles is set to 10% the maximum velocity.

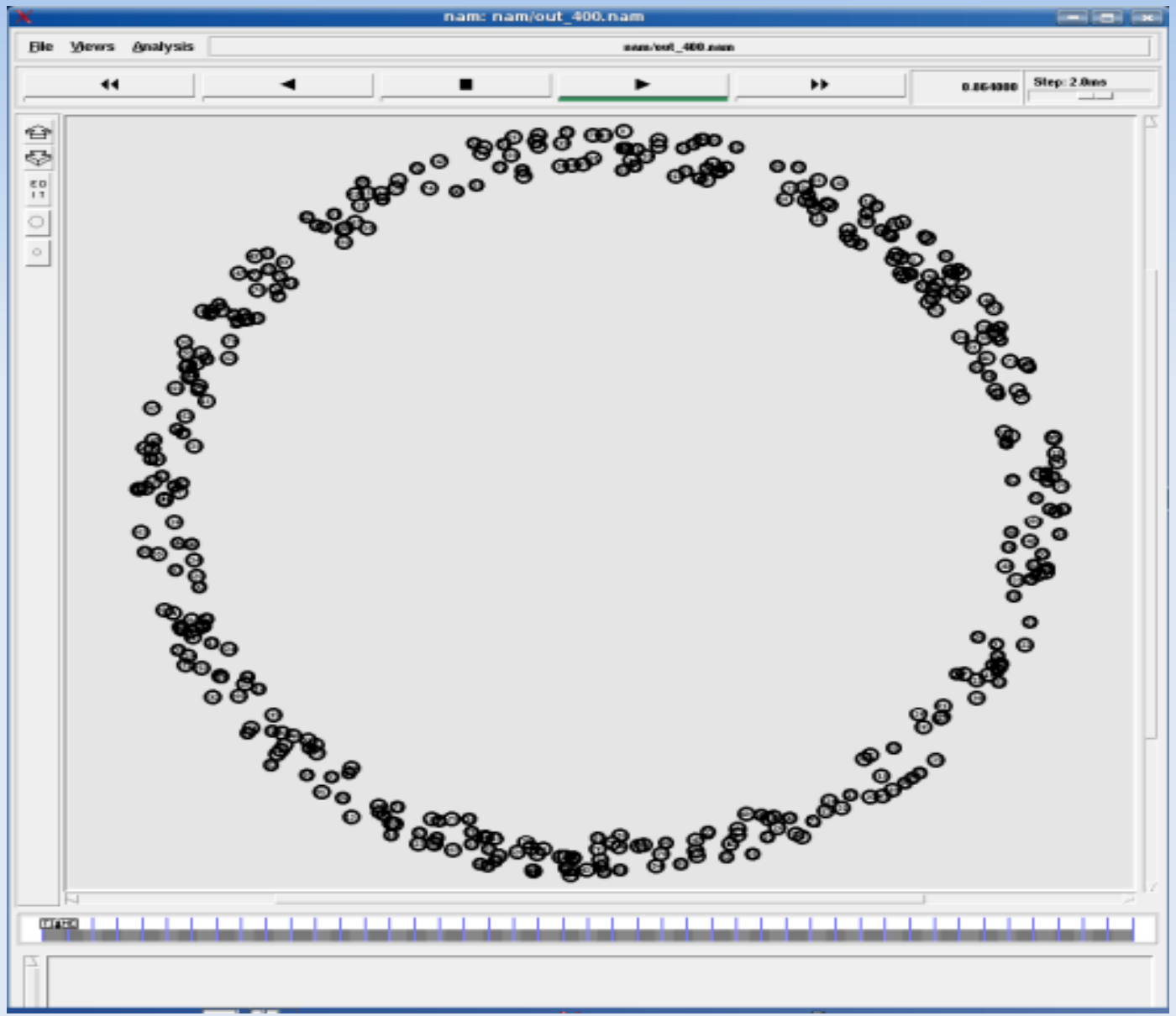
Network Traffic

- AC[0] is used to transmit emergency warnings. Nodes are randomly selected to transmit an emergency warning and they are transmitted once every 100 *ms* for 1.5 *s* with 10% jitter.
- AC[2] is used to transmit a vehicle's state every 100 *ms* seconds with 10% jitter with .
- Approximately 20% of the traffic is AC[0] and 80% of the traffic is AC[2].

Metrics Used to Evaluate Protocols

- Channel access time: the elapsed from when a packets is passed to the MAC layer until it is put on the channel.
- Reception rate: the percentage of packets successfully received at a distance d± 5m from the sender.

Simulation Topology



Simulation 400 Nodes

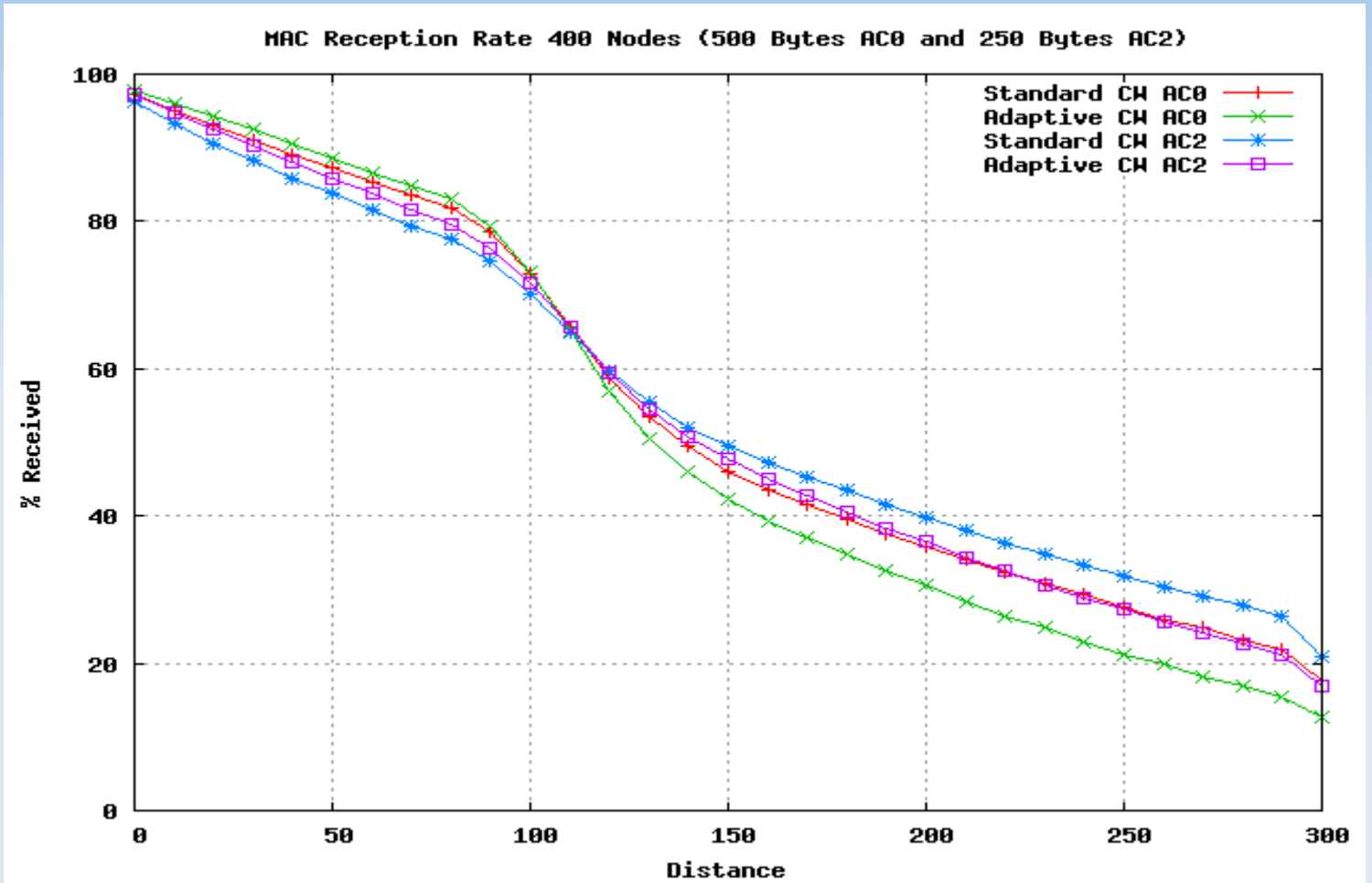


Table 4.7: Average Access Delay for 400 Nodes

	Standard CW	Modified CW
AC0	0.001018	0.001580
AC2	0.001861	0.010133

Simulations 600 Nodes

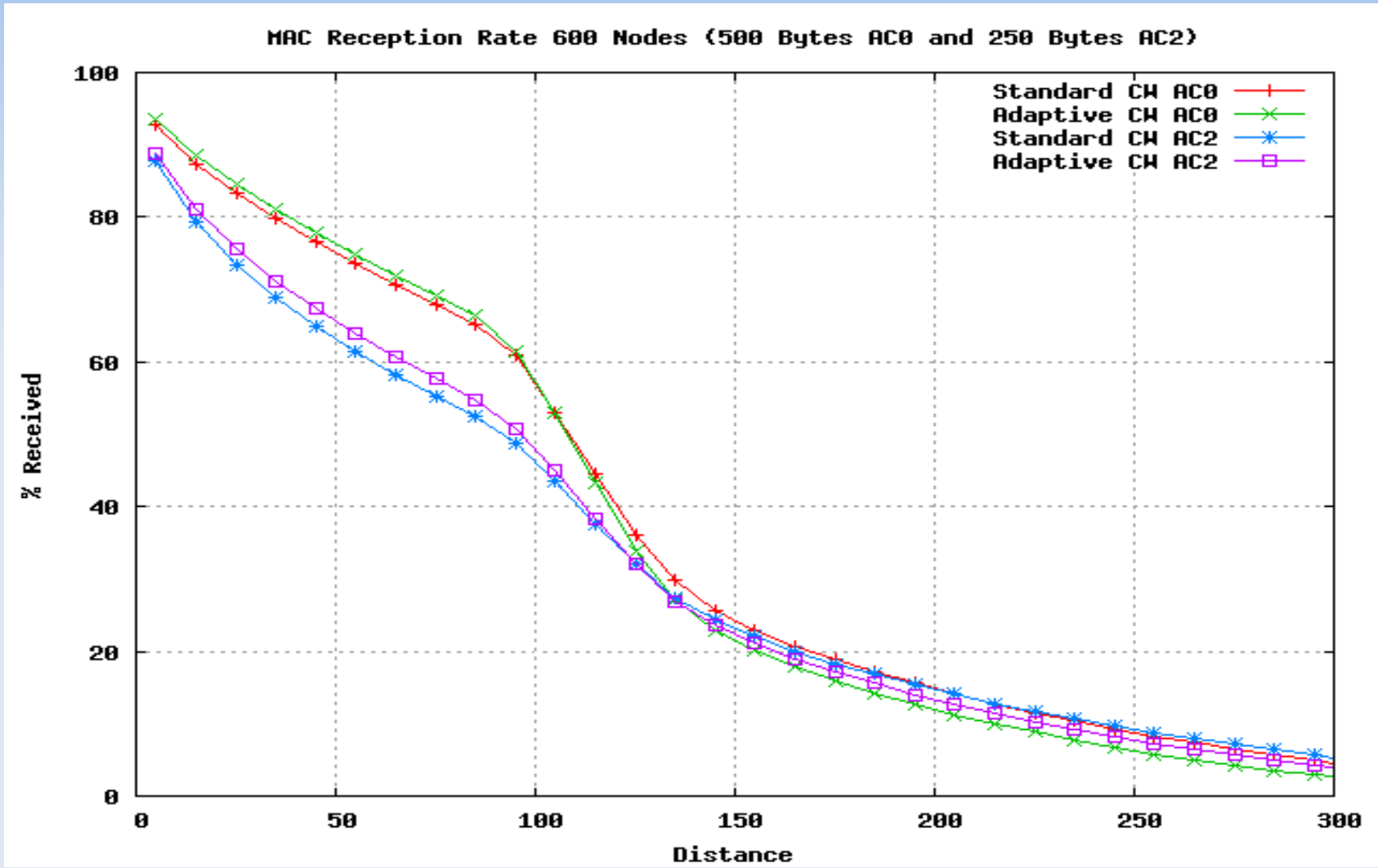


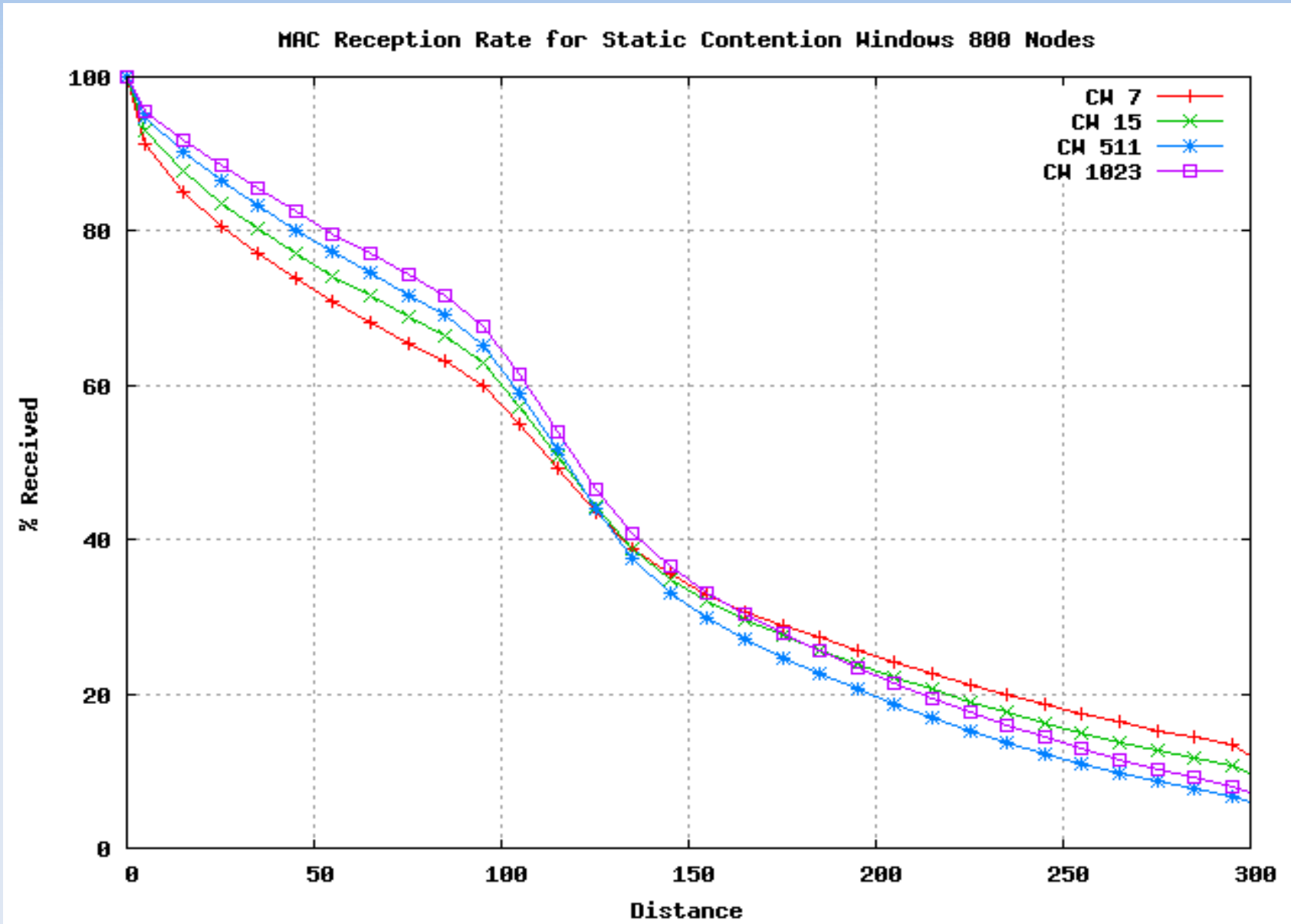
Table 4.8: Average Access Delay for 600 Nodes

	Standard CW	Modified CW
AC0	0.002428	0.001580
AC2	0.010687	0.066329

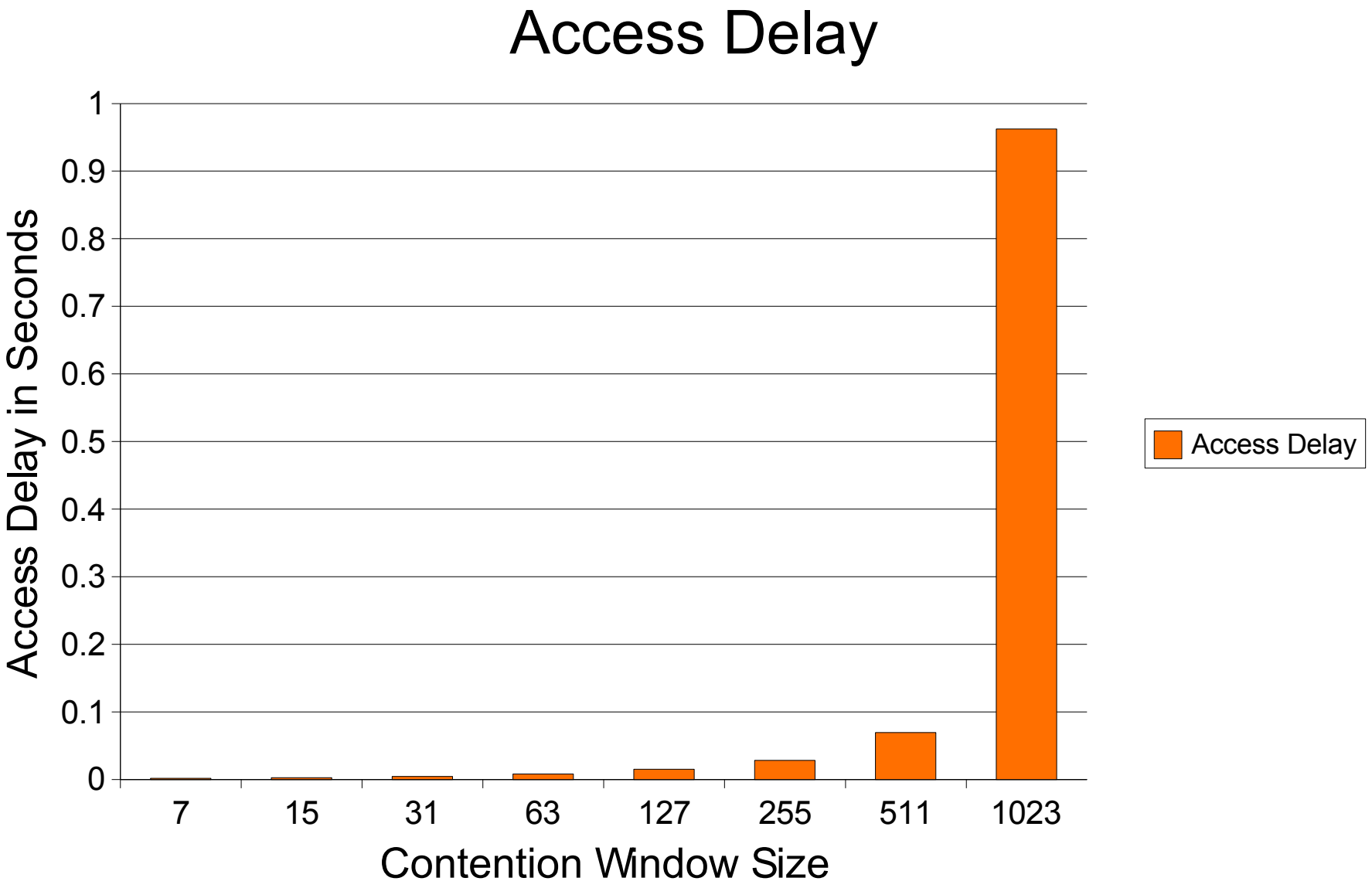
The Effect Static CW Sizes Have on Reception

- The results of the simulations for the modified CW algorithm show the CW algorithm slightly increased the reception rate.
- For this reason, a set of simulations was used to determine the effect that different values of CW_{min} have on the reception rate.
- The results of the additional simulations show that it is unlikely to be able to increase the reception rate by more than 5%-10%.
- One interesting finding was the a larger CW results in a lower probability of reception at large distance, while the reception is improved at short distances with large windows.
- For light network traffic the reception rate does not improved much as a result of varying the CW.
- For heavy traffic the rate improves but the queues can also fill up causing packets to be dropped and causing excessive delays.

Fixed Contention Windows



Access Delay



Dynamic Contention Window Adjustment

- Dynamically adjusting the CW will have a some impact on the reception rate.
- The previous results showed that a 5% to 6% increase can be expected by increasing the CW from CW_{min} = 7 to CW_{min} = 511 at certain distances.
- The dynamic CW would likely result in even a lower improvement in the reception rate than this.

Conclusion

- Dynamically adjusting the contention window improves the rate of reception under certain conditions.
- If the amount of network traffic is close to the theoretical limit, it would be beneficial instead to drop some packets or adjust the rate that packets are transmitted at. In this case increasing the CW has a negligible impact.

Future Work

- Adaptive Transmission Rate
- Dynamic Power Transmission Control

Adaptive Transmission Rate

- Due to the hidden terminal problem it is unrealistic to expect to achieve anywhere near 100% delivery of broadcast frames.
- In the case of a highly loaded network, increasing the CW to 1023 slots will result in unacceptable access delay.
- In the case of a highly load network nodes will have to decrease their transmission rate.
- The feedback from the network can be used to reduce the transmission rate of low priority traffic.

Dynamic Power Transmission Control

- Controlling the communication range by adjusting the transmission power can be used to mitigate the adverse effects of highly dense nodes.
- The choice of the communication range has a direct impact on network connectivity.
- Most approaches use the node's density to estimate the transmission range.
- The transmission range can also be adjusted to keep the network load on the channel below a certain threshold using this method.