A Coherent Theory of Random-Access Networks

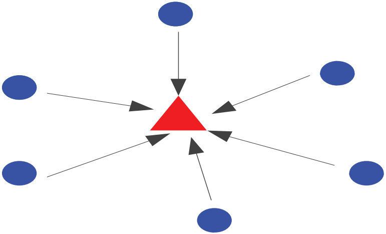
Lin Dai

Department of Electronic Engineering
City University of Hong Kong

lindai@cityu.edu.hk

April, 2017

Multiple Access



Multiple nodes transmit to a common receiver: How to share the channel?

- Centralized Access: A central controller performs resource allocation/optimization.
- Random Access: Each node determines when/how to access in a distributed manner.

Applications of Random Access

- WiFi networks
- Cellular networks
- Sensor networks
- Machine-to-machine (M2M) communications, Vehicle-to-vehicle (V2V) communications, Internet-of-things (IoT),......

Design of Random-Access Networks: Three Key Questions

For each node:

- When to start a transmission?
- When to end a transmission?
- What if the transmission fails?

Question 1: When to Start a Transmission?

- Transmit if packets are awaiting in the queue.
  — Aloha [Abramson'1970]
- A more "polite" solution: Transmit if packets are awaiting in the queue **and the channel is sensed idle**.
  — Carrier Sense Multiple Access (CSMA) [Kleinrock&Tobagi'1975]

Question 2: When to End a Transmission?

- Stop when a packet transmission is completed.
- Any "smarter" solution?
  - — Stop when other on-going transmissions are sensed.
  - — What if nodes cannot sense the channel while transmitting?

Question 3: What if the Transmission Fails?

The definition of transmission failure depends on what type of receivers is adopted. Various assumptions on the receiver have been made, which can be broadly divided into three categories.

- *Collision*: When more than one node transmit their packets simultaneously, a collision occurs and none of them can be successfully decoded. A packet transmission is successful only if there are no concurrent transmissions.
- *Capture*: Each node's packet is decoded independently by treating others' as background noise. A packet can be successfully decoded as long as its received signal-to-interference-plus-noise ratio (SINR) is above a certain threshold.
- *Joint-decoding*: Multiple nodes' packets are jointly decoded.

Question 3: What if the Transmission Fails?

Backoff if the transmission fails.

- Probability-based: Retransmit with a certain probability at each time slot.
- Window-based: Choose a random value from a window and count down. Retransmit when the counter is zero.

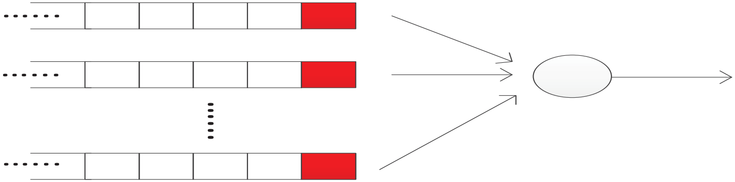
In general, backoff can be characterized as a sequence of transmission probabilities $\{q_t\}$, where q_t denotes the transmission probability of the head-of-line (HOL) packet in each node's buffer at time slot t.

It is usually assumed that the transmission probability q_t is adjusted according to the number of collisions that the HOL packet has experienced by time slot t, i.e., $q_t = \mathcal{Q}(i)$, where $\mathcal{Q}(i)$ is an arbitrary monotonic non-increasing function of the number of collisions i. With Exponential Backoff [Metcalfe&Boggs'1976], for instance, $\mathcal{Q}(i) = 2^{-i}$.

Performance Analysis of Random-Access Networks: Three Key Questions

- How to model a random-access network?
- How to evaluate the performance of a random-access network?
- How to optimize the performance of a random-access network?

Question 1: How to Model a Random-Access Network?



- A random-access network can be regarded as a multi-queue-single-server system.
- Numerous models have been proposed, which can be broadly divided into two categories: channel-centric and node-centric.

Question 1: How to Model a Random-Access Network?

- Channel-centric modeling: to characterize the aggregate traffic of all the nodes.

  — [Abramson'1970], [Kleinrock&Tobagi'1975], ...: The aggregate traffic is modeled as a Poisson random variable.

  — The aggregate traffic is determined by the input traffic and backoff mechanism of each node, which may not always be approximated as a Poisson random variable.

- Node-centric modeling: to characterize the queueing behavior of each node.

  — [Tsybakov'1979]: A two-node buffered Aloha network is modeled as a 2-dimensional random walk.

  — [Rao&Ephremides'1988], [Anantharam'1991], ...: Generalization to an n-node buffered Aloha network.

  — Modeling complexity quickly increases with the number of nodes.

Question 1: How to Model a Random-Access Network?

To establish a scalable node-centric model [Dai'12], [Dai'13]:

- Treat each node's queue as an independent queueing system with identically distributed service time.

- Characterization of the service time distribution includes two parts:

  — State characterization of each individual head-of-line (HOL) packet;

  — Fixed-point equations of steady-state probability of successful transmission of HOL packets.

Question 2: How to Evaluate the Performance of a Random-access Network?

- Network Throughput: the average number of successful transmitted packets of the network per time slot.
- Delay
  1) Queueing delay (waiting time): the time interval from the packet's arrival to the instant that it becomes the HOL packet;
  2) Access delay (service time): the time interval from the instant that it becomes the HOL packet to its successful transmission.
- Stability
  1) The network is stable if the network throughput is equal to the aggregate input rate.
  2) The network is stable if the mean access/queueing delay is finite.
  ...
- Network Sum Rate: the average number of successfully transmitted information bits of the network per time slot.

Question 2: How to Evaluate the Performance of a Random-access Network?

To establish a unified theoretical framework within which effects of key system parameters on various performance metrics can be evaluated in a systematic manner:

	Aloha	CSMA	802.11 DCF
Network Throughput	[Dai'12]	[Dai'13], [Sun-Dai'16]	[Dai-Sun'13], [Gao-Sun-Dai'13], [Gao-Dai'13], [Gao-Sun-Dai'14], [Sun-Dai'15], [Sun-Dai'16]
Delay	[Dai'12]	[Dai'13], [Sun-Dai'16]	[Dai-Sun'13], [Sun-Dai'15], [Sun-Dai'16]
Stability	[Dai'12]	[Dai'13]	[Dai-Sun'13]
Network Sum Rate	[Li-Dai'16]	[Sun-Dai'17]	

Question 3: How to Optimize the Performance of a Random-access Network?

- Performance of random-access networks crucially depends on backoff parameters.
- With fixed backoff parameters, a random-access network suffers from significant performance degradation as the network size or the traffic input rates increase.

For given network size and traffic input rate of each node, how to properly set the backoff parameters to optimize the network performance?

Question 3: How to Optimize the Performance of a Random-access Network?

Network Throughput Optimization:

- Aloha: [Dai'12]
- CSMA: [Dai'13] $\xrightarrow{\text{Finite Retry Limit}}$ [Sun-Dai'16]
- 802.11 DCF: [Dai-Sun'13]

  $\xrightarrow{\text{Heterogeneous}}$ [Gao-Sun-Dai'13], [Gao-Dai'13], [Gao-Sun-Dai'14]

  $\xrightarrow{\text{General Backoff}}$ [Sun-Dai'15]

  $\xrightarrow{\text{Finite Retry Limit}}$ [Sun-Dai'16]

Question 3: How to Optimize the Performance of a Random-access Network?

Delay Optimization:

- Aloha: [Dai'12]
- CSMA: [Dai'13] $\xrightarrow{\text{Finite Retry Limit}}$ [Sun-Dai'16]
- 802.11 DCF: [Dai-Sun'13]

  $\xrightarrow{\text{General Backoff}}$ [Sun-Dai'15]

  $\xrightarrow{\text{Finite Retry Limit}}$ [Sun-Dai'16]

Question 3: How to Optimize the Performance of a Random-access Network?

Network Sum Rate Optimization:

- Aloha: [Li-Dai'16]
- CSMA: [Sun-Dai'17]

References

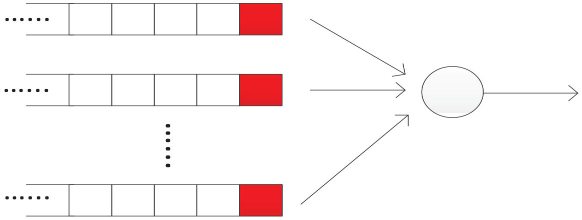
- L. Dai, "Stability and delay analysis of buffered Aloha networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 8, pp. 2707–2719, Aug. 2012.
- L. Dai, "Toward a coherent theory of CSMA and Aloha," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3428–3444, Jul. 2013.
- L. Dai and X. Sun, "A unified analysis of IEEE 802.11 DCF networks: stability, throughput and delay," *IEEE Trans. Mobile Computing*, vol. 12, no. 8, pp. 1558–1572, Aug. 2013.
- Y. Gao, X. Sun, and L. Dai, "Throughput Optimization of Heterogeneous IEEE 802.11 DCF Networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 398–411, Jan. 2013.
- Y. Gao and L. Dai, "Optimal Downlink/Uplink Throughput Allocation for IEEE 802.11 DCF Networks," *IEEE Wireless Commun. Letters*, vol. 2, no. 6, pp. 627–630, Dec. 2013.
- Y. Gao, X. Sun, and L. Dai, "IEEE 802.11e EDCA Networks: Modeling, Differentiation and Optimization," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3863–3879, July 2014.

References

- X. Sun and L. Dai, "Backoff Design for IEEE 802.11 DCF Networks: Fundamental Tradeoff and Design Criterion," *IEEE/ACM Trans. Networking*, vol. 23, no. 1, pp. 300–316, Feb. 2015.
- Y. Li and L. Dai, "Maximum Sum Rate of Slotted Aloha with Capture," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 690–705, Feb. 2016.
- X. Sun and L. Dai, "Performance Optimization of CSMA Networks with A Finite Retry Limit," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 5947–5962, Sept. 2016.
- X. Sun and L. Dai, "Fairness-constrained Maximum Sum Rate of Multi-rate CSMA Networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1741–1754, Mar. 2017.

Modeling of Random-Access Networks

Node-Centric Modeling



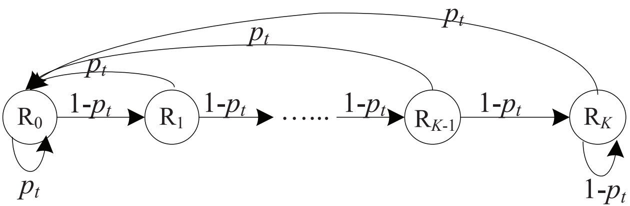
- An n-node buffered random-access network can be modeled as an n-queue-single-server system.
- Key assumption: Each node's queue can be regarded as an independent queueing system when the number of nodes n is large.
- Key steps: Characterization of each head-of-line (HOL) packet's state transition and steady-state probability of successful transmission of HOL packets.

State Characterization of HOL Packet

The state transition of each HOL packet can be modeled as a discrete-time Markov renewal process $(\mathbf{X}^h, \mathbf{V}^h) = \{(X_j^h, V_j^h), j = 0, 1, \ldots\}$.

The embedded Markov chain $\mathbf{X}^h = \{X_j^h\}$ is:

Without Sensing



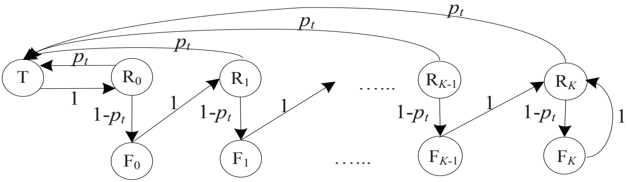
- K: cutoff phase. The HOL packet stays at State R_K if the number of collisions exceeds K.
- p_t: probability of successful transmission of HOL packets at time slot t. $\lim_{t\to\infty} p_t = p$.

State Characterization of HOL Packet

The state transition of each HOL packet can be modeled as a discrete-time Markov renewal process $(\mathbf{X}^h, \mathbf{V}^h) = \{(X_j^h, V_j^h), j = 0, 1, \ldots\}$.

The embedded Markov chain $\mathbf{X}^h = \{X_j^h\}$ is:

With Sensing



- The states of $\{X_j^h\}$ are divided into three categories: 1) waiting to request (State R_i, $i = 0, \ldots, K$), 2) collision (State F_i, $i = 0, \ldots, K$) and 3) successful transmission (State T).
- p_t: probability of successful transmission of HOL packets at mini-slot t given that the channel is idle at $t-1$. $\lim_{t\to\infty} p_t = p$.

Dynamic Trajectory of p_t

- For each HOL packet, its transmission is successful if and only if all the other $n-1$ nodes are either idle with an empty queue, or, with a State-R_i HOL packet but not requesting any transmission.

Without Sensing

$$p_{t+1} = \left\{1 - \rho_t + \sum_{i=1}^{K} \rho_t \tilde{\pi}_{R_i,t}(1-q_i)\right\}^{n-1}.$$

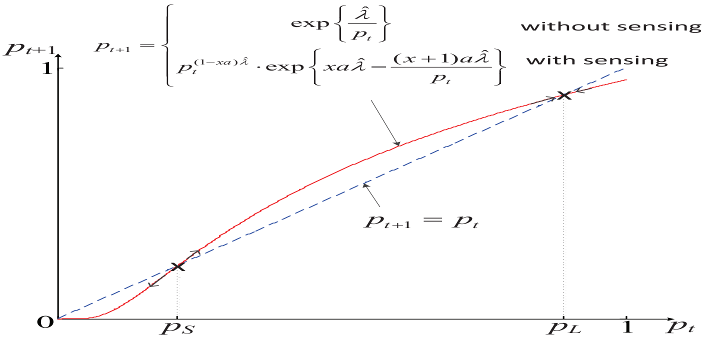
With Sensing

$$p_{t+1} = \left\{\frac{1 - \rho_t + \rho_t \cdot \sum_{i=0}^{K} \tilde{\pi}_{R_i,t} \cdot (1-q_i)}{1 - \rho_t + \rho_t \cdot \sum_{i=0}^{K} \tilde{\pi}_{R_i,t}}\right\}^{n-1}.$$

q_i: Transmission probability of a State-R_i HOL packet.
ρ_t: Offered load at time slot t.
$\tilde{\pi}_{i,t}$: Probability that the HOL packet stays at State R_i at time slot t.

- Approximation: $\rho_t \approx \rho(p_t)$ and $\tilde{\pi}_{R_i,t} \approx \tilde{\pi}_{R_i}(p_t)$, $i = 0, \ldots, K$.

Bi-stable Property



- $\hat{\lambda}$: aggregate input rate of nodes.
- a: time-slot length (with sensing). a is determined by the ratio of the propagation delay to the packet length.
- x: collision-detection time in the unit of time slots, i.e., it takes each node x time slots to detect the transmission failure and abort the transmission.

Bi-stable Property

- If $p_t \geq p_S$ at any t and $\lim_{t\to\infty} \rho_t = \rho \leq 1$: $\lim_{t\to\infty} p_t \to p_L$.

  p_L and p_S are roots of the following fixed-point equation:

$$p = \begin{cases} \exp\left\{\frac{\hat{\lambda}}{p}\right\} & \text{without sensing} \\ \exp\left\{\frac{xa\hat{\lambda}}{1-(1-xa)\hat{\lambda}}\right\} \cdot \exp\left\{-\frac{(x+1)a\hat{\lambda}}{1-(1-xa)\hat{\lambda}} \cdot \frac{1}{p}\right\} & \text{with sensing.} \end{cases}$$

- Otherwise, the network becomes saturated and all the nodes are busy with non-empty queues. In this case, $\lim_{t\to\infty} p_t \to p_A$, where p_A is the single non-zero root of

$$p = \exp\left\{-\frac{n}{\sum_{i=0}^{K-1} \frac{p(1-p)^i}{q_i} + \frac{(1-p)^K}{q_K}}\right\},$$

  if the transmission probability $\{q_i\}$ is a monotonic non-increasing sequence.

Desired Stable Point p_L and Undesired Stable Point p_A

- Desired Stable Point p_L is determined by the aggregate input rate $\hat{\lambda}$ and independent of the backoff parameters $\{q_i\}$.
- Undesired Stable Point p_A is determined by the backoff parameters $\{q_i\}$ and the number of nodes n.

	Without Sensing	With Sensing
Desired Stable Point p_L	$p_L^{Aloha} = \exp\left\{\mathbb{W}_0(-\hat{\lambda})\right\}$	$p_L^{CSMA} = \exp\left\{\mathbb{W}_0\left(-\frac{(x+1)a\hat{\lambda}}{1-(1-xa)\hat{\lambda}} \cdot \exp\left\{-\frac{xa\hat{\lambda}}{1-(1-xa)\hat{\lambda}}\right\}\right) + \frac{xa\hat{\lambda}}{1-(1-xa)\hat{\lambda}}\right\}$
Undesired Stable Point p_A	$p_A = \exp\left\{-\frac{n}{\sum_{i=0}^{K-1}\frac{p_A(1-p_A)^i}{q_i} + \frac{(1-p_A)^K}{q_K}}\right\}$	

Service Rate and Service Time Distribution

- Service rate:

  Without Sensing: $-p \ln p$.

  With Sensing: $\frac{-p \ln p}{(x+1)a-(1-xa)p \ln p-xap}$.

- Service time distribution:

  Without Sensing: $G_{D_i}(z) = pG_{Y_i}(z) + (1-p)G_{D_{i+1}}(z)$, $i = 0, \ldots, K-1$, and $G_{D_K}(z) = G_{Y_K}(z)$.

  With Sensing: $G_{D_i}(z) = pzG_{Y_i}(z) + (1-p)z^{xa}G_{Y_i}(z)G_{D_{i+1}}(z)$, $i = 0, \ldots, K-1$, and $G_{D_K}(z) = pzG_{Y_K}(z) + (1-p)z^{xa}G_{Y_K}(z)G_{D_K}(z)$.

  $G_X(z)$ denotes the probability generating function of X. D_i denotes the time spent from the beginning of State R_i until the service completion, and Y_i denotes the holding time of a HOL packet in State R_i, $i = 0, \ldots, K$.

Summary

- The key to node-centric modeling lies in proper characterization of 1) the state transition process of each HOL packet, and 2) the steady-state probability of successful transmission of HOL packets.
- Based on the proposed unified analytical framework, effects of key parameters on a wide range of performance metrics such as network throughput, access delay, sum rate and stability, of various random-access networks can be evaluated in a systematic manner.
- The proposed analytical framework can further facilitate performance optimization to reveal the fundamental limits of random-access networks, and to show how to properly set the key system parameters to achieve the limiting performance.

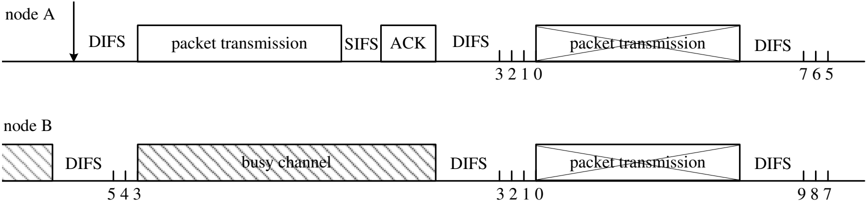
Application to WiFi (IEEE 802.11) Networks

Distributed Coordination Function (DCF) of IEEE 802.11 Networks

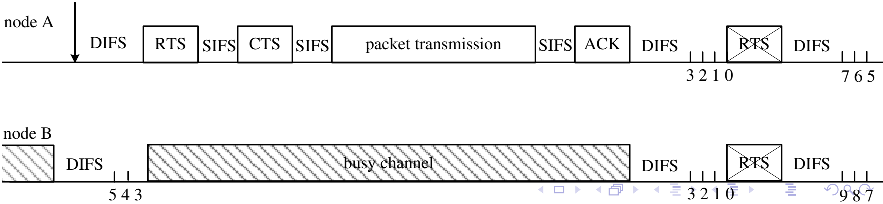
- Carrier Sense Multiple Access (CSMA) with Binary Exponential Backoff (BEB).
- Two access mechanisms: basic access and request-to-send/clear-to-send (RTS/CTS) access.

Basic Access and RTS/CTS Access

- Basic Access:



- RTS/CTS Access:

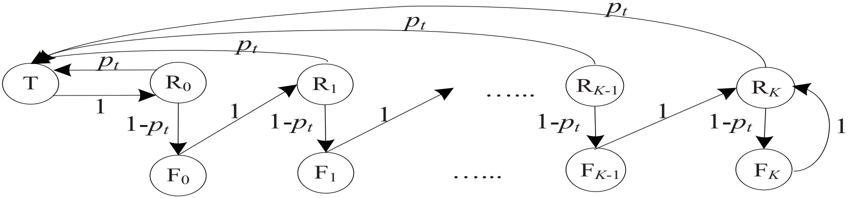


Backoff

- Choose a random value from $\{0, ..., W_i - 1\}$, where W_i is the backoff window size after the i-th collision, $i = 0, 1, ...$
- Count down when the channel is idle.
- Binary Exponential Backoff (BEB): $W_i = W \cdot 2^{\min(i,K)}$, where K is the cutoff phase.

Node-Centric Modeling

- State characterization of Head-of-Line (HOL) packet



- Characterization of network steady-state points based on the fixed-point equations of limiting probability of successful transmission of HOL packets $p = \lim_{t\to\infty} p_t$.

State Characterization of HOL Packets: State T

- State T: The HOL packet makes a successful transmission.
- Mean holding time τ_T (in unit of time slots) at State T depends on the access mechanism.

- Basic Access:

$$\tau_T^{Basic} = \frac{\text{PHY+MAC+Payload+SIFS+ACK+DIFS}}{\text{slot time}}$$

- RTS/CTS Access:

$$\tau_T^{RTS} = \frac{\text{RTS+SIFS+CTS+SIFS+PHY+MAC+Payload+SIFS+ACK+DIFS}}{\text{slot time}}$$

State Characterization of HOL Packets: State F_i

- State F_i: The HOL packet experiences a collision.
- Mean holding time τ_F (in unit of time slots) at State F_i depends on the access mechanism.

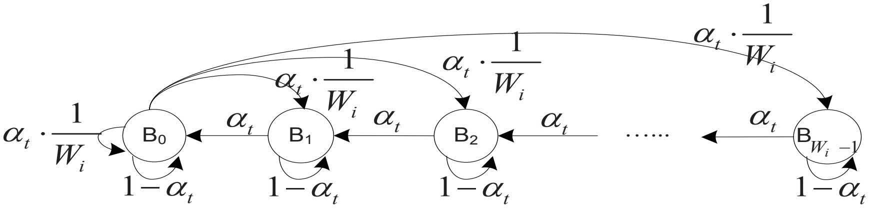
- Basic Access:

$$\tau_F^{Basic} = \frac{\text{PHY+MAC+Payload+DIFS}}{\text{slot time}}$$

- RTS/CTS Access:

$$\tau_F^{RTS} = \frac{\text{RTS+DIFS}}{\text{slot time}}$$

- State R_i: The HOL packet waits to request a transmission.
- Mean holding time τ_{R_i} (in unit of time slots) at State R_i depends on the backoff window size W_i.



- $\tau_{R_i} = \frac{1}{2\alpha} \cdot (1 + W_i)$, where $\alpha = \lim_{t\to\infty} \alpha_t$ is the limiting probability of sensing the channel idle, which is given by $\alpha = \frac{1}{1+\tau_F-\tau_F p-(\tau_T-\tau_F)p \ln p}$.

Network Steady-State Points

Desired stable point p_L

$$p_L = \exp\left\{ \mathbb{W}_0\left(-\frac{\hat{\lambda}(1+\tau_F)/\tau_T}{1-(1-\tau_F/\tau_T)\hat{\lambda}} \cdot \exp\left\{ -\frac{\hat{\lambda}\tau_F/\tau_T}{1-(1-\tau_F/\tau_T)\hat{\lambda}} \right\} \right) + \frac{\hat{\lambda}\tau_F/\tau_T}{1-(1-\tau_F/\tau_T)\hat{\lambda}} \right\}.$$

- p_L is determined by the aggregate input rate $\hat{\lambda}$, and the holding time in successful transmission and collision states, τ_T and τ_F.

Undesired stable point p_A

- When the network becomes saturated, it shifts to the undesired stable point p_A, which is the root of the following fixed-point equation:

$$p = \exp\left\{ -\frac{2n}{W\left(\frac{p}{2p-1} + \left(1-\frac{p}{2p-1}\right)(2(1-p))^K \right)} \right\}.$$

- p_A is determined by the network size n and the backoff parameters, i.e., initial backoff window size W and the cutoff phase K.

$\mathbb{W}_0(.)$ is the principal branch of the Lambert W function.

Maximum Network Throughput

- The maximum network throughput $\hat{\lambda}_{\max}$ of IEEE 802.11 DCF networks is

$$\hat{\lambda}_{\max} = \frac{-\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right)}{\tau_F/\tau_T - (1-\tau_F/\tau_T)\,\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right)}.$$

- The optimal initial backoff widow size W_m for achieving $\hat{\lambda}_{\max}$ is

$$W_m = n \cdot \frac{-\frac{4(1+\tau_F)}{\tau_F}\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right) - 2}{\frac{1+\tau_F}{\tau_F}\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right)\ln\left(-\frac{1+\tau_F}{\tau_F}\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right)\right)}.$$

To achieve the maximum network throughput, the initial backoff window size should be linearly increased with the network size n.

Maximum Network Throughput: Basic Access versus RTS/CTS Access

Table : Parameter Setting (802.11n)

Packet payload	4096B
MAC header	36B
PHY header	20μs
ACK	14B+PHY header
RTS	20B+PHY header
CTS	14B+PHY header
Slot Time	9μs
SIFS	16μs
DIFS	34μs
Basic Rate	6 Mbps
Data Rate	57.8 Mbps

Basic Access

- $\tau_T^{Basic} = 75.7$ time slots and $\tau_F^{Basic} = 69.5$ time slots.
- $\hat{\lambda}_{\max}^{Basic} = 0.85$.
- $W_m^{Basic} = 10.4n$.

RTS/CTS Access

- $\tau_T^{RTS} = 88.7$ time slots and $\tau_F^{RTS} = 9$ time slots.
- $\hat{\lambda}_{\max}^{RTS} = 0.94$.
- $W_m^{RTS} = 2.7n$.

Access Delay

- D_0: Service time of HOL packets (which is also the access delay of packets).
- Based on the state transition of HOL packets, we can obtain the probability generating function $G_{D_0}(z)$, from which the n-th moment of D_0 can be derived, $n = 1, 2, ...$

Access Delay

Mean Access Delay:

$$E[D_0]=\tau_T+\frac{1-p}{p}\tau_F+\frac{1+\tau_F-\tau_F p-(\tau_T-\tau_F)p\ln p}{2}$$
$$\cdot\left(\sum_{i=0}^{K-1}(1-p)^i W_i+\frac{(1-p)^K}{p}W_K+\frac{1}{p}\right).$$

Second Moment of Access Delay:

$$E[D_0^2]=\frac{3+p-3\alpha p}{6\alpha^2 p^2}+\left(\frac{2}{\alpha}+(2-p)\tau_F\right)\frac{(1-p)\tau_F}{p^2}+\left(\tau_T+\frac{2(1-p)}{p}\tau_F+\frac{1}{\alpha p}\right)\tau_T$$
$$+\left(\frac{1-\alpha}{2\alpha^2}+\frac{1}{2\alpha^2 p}+\frac{p\tau_T+(1-p)\tau_F}{\alpha p}\right)\sum_{i=0}^{\infty}(1-p)^i W_i+\frac{1}{\alpha}\left(\tau_F+\frac{1}{2\alpha}\right)\cdot\sum_{i=0}^{\infty}$$
$$\left(\sum_{j=i+1}^{\infty}(1-p)^j W_j\right)+\frac{1}{3\alpha^2}\sum_{i=0}^{\infty}(1-p)^i W_i^2+\frac{1}{2\alpha^2}\sum_{i=0}^{\infty}W_i\left(\sum_{j=i+1}^{\infty}(1-p)^j W_j\right).$$

Mean Access Delay with BEB

At the desired stable point $p_L \approx 1$

$$E[D_{0,p=p_L}] = \tau_T + \frac{1+W}{2}.$$

At the undesired stable point p_A

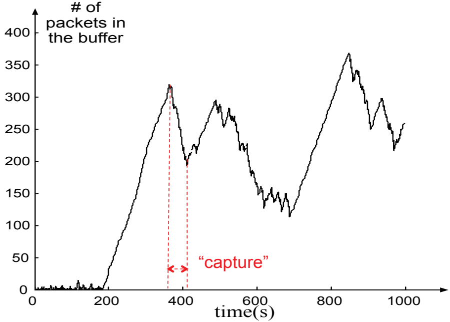
$$E[D_{0,p=p_A}] = \left(\tau_T + \tau_F \frac{1-p_A}{p_A} + (1+\tau_F(1-p_A) - (1-\tau_F)\, p_A \ln p_A)\right.$$
$$\left.\cdot\left(\frac{1}{2p_A} + \frac{W}{2}\left(\frac{1}{2p_A-1} + \left(\frac{1}{p_A} - \frac{1}{2p_A-1}\right)(2(1-p_A))^K\right)\right)\right).$$

- $E[D_{0,p=p_A}]$ is inversely proportional to the network throughput.
- $\min_W E[D_{0,p=p_A}] = \frac{n}{\hat{\lambda}_{\max}}$, which is achieved when $W = W_m$.

Second Moment of Access Delay with BEB

At the undesired stable point p_A

$$E[D_{0,p=p_A}^{2,K=\infty}] = \infty \text{ when } W \leq \frac{4n}{3\ln\frac{4}{3}} \approx 4.63n.$$



- Capture phenomenon: Some node captures the channel and produces a continuous stream of packets, while others have to wait for a long time.
  - Poor queueing delay performance
  - Short-term unfairness

Effect of Backoff

In general, a backoff scheme can be characterized by a sequence of backoff window sizes $\{W_i\}$.

With a monotonic increasing sequence of backoff window sizes $\{W_i\}$, the limiting growth rate of backoff window size is

$$\lim_{i\to\infty} \frac{W_{i+1}}{W_i} \geq 1.$$

Accordingly, we can divide the backoff schemes into two categories:

- Aggressive Backoff: $\lim_{i\to\infty} \frac{W_{i+1}}{W_i} > 1$.
- Mild Backoff: $\lim_{i\to\infty} \frac{W_{i+1}}{W_i} = 1$.

Tradeoff between Throughput and Delay

Theorem

The maximum network throughput $\hat{\lambda}_{\max}$ *is given by*

$$\hat{\lambda}_{\max} = \frac{-\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right)}{\tau_F/\tau_T - (1-\tau_F/\tau_T)\mathbb{W}_0\left(-\frac{1}{e(1+1/\tau_F)}\right)}.$$

Corollary

$\hat{\lambda}_{\max}$ *is achievable as* $n \to \infty$ *if and only if* $K = \infty$ *and* $\lim_{i\to\infty} \frac{W_{i+1}}{W_i} > 1$.

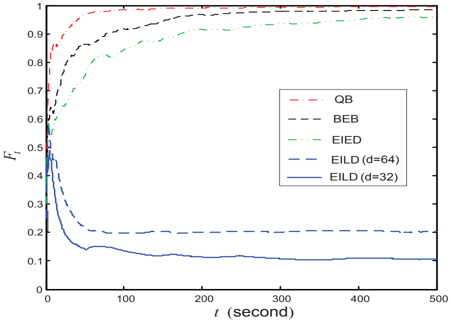
Theorem

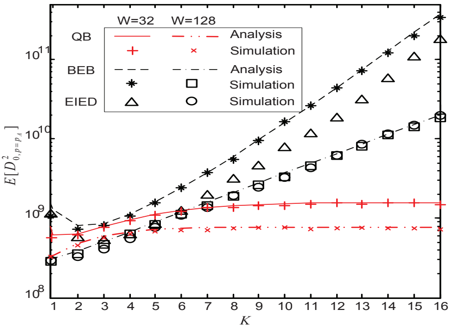
If $\lim_{i\to\infty} \frac{W_{i+1}}{W_i} = 1$, $E[D^{2,K=\infty}_{0,p=p_A}] < \infty$ *for any finite* $n < \infty$*; otherwise, if* $\lim_{i\to\infty} \frac{W_{i+1}}{W_i} > 1$, *there exists* $n' < \infty$ *such that* $E[D^{2,K=\infty}_{0,p=p_A}] = \infty$ *when* $n > n'$.

Tradeoff between Throughput and Delay

- Both aggressive backoff and mild backoff have the same maximum network throughput $\hat{\lambda}_{\max}$, which is solely determined by the access mechanism, i.e., the basic access or the RTS/CTS access.
- With aggressive backoff schemes, the maximum network throughput can be achieved even as the network size n goes to infinity. But an infinite second moment of access delay may be incurred if the network size exceeds a certain threshold.
- Mild backoff schemes can reach a better balance: The maximum network throughput and a bounded second moment of access delay can both be achieved as long as the network size is finite.

Comparison of BEB and QB: Fairness and Second Moment of Access Delay





Jain's Fairness Index F_t versus time t in saturated IEEE 802.11 DCF networks with BEB, QB, EIED and EILD. $W = 32$, $K = 6$ and $n = 50$.

Second moment of access delay $E[D^2_{0,p=p_A}]$ versus cutoff phase K in saturated IEEE 802.11 DCF networks with BEB, QB and EIED. $n = 50$.

Further Extensions

- From homogeneous to heterogeneous:
  - Distinct traffic input rates: [Gao-Sun-Dai'13]
  - Distinct priority levels: [Gao-Sun-Dai'14]
  - Distinct data transmission rates: [Sun-Dai'17]
- From infinite retry limit to finite retry limit: [Sun-Dai'16]

Ongoing and Future Work

- Delay-constrained maximum sum rate
- More sophisticated receivers
- Multiple basic service sets (BSSs) and multiple co-existing standards

The End

Thank You!