AI/ML in Networking

Lecture 2: Wi-Fi Resource Allocation

Francesc Wilhelmi (francisco.wilhelmi@upf.edu)

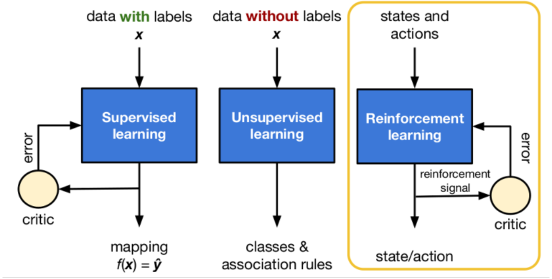
Università degli Studi di Cagliari

Outline

Types of Machine Learning algorithms

Types of machine learning:

- **Supervised learning (SL):** use labeled examples to drive the adjustment of the model.
- **Unsupervised learning (UL):** find patterns in the data without using labels (related to statistics).
- **Reinforcement learning (RL):** learning from experience (agent-based).



Types of tasks

Classification: assign data points to predefined categories or classes.

- Binary Classification: Predicting one of two classes (e.g., yes/no).
- Multi-class Classification: Predicting one of more than two classes (e.g., types of animals).
- Multi-label Classification: Assigning multiple labels to each data point (e.g., movie genres).

Regression: The goal is to predict a continuous numerical value.

- Linear regression: Assume a linear relationship between the variables.
- Polynomial regression: Assume a polynomial relationship between the variables.
- Non-linear Regression: Assume a non-linear relationship between the variables.

Classification

y = {yellow, gray}
y = {churn, no churn}
y = {increase, unchanged, decrease}
y = {blonde, gray, brown, red, black}
y = {job 1, job 2, ... , job n}

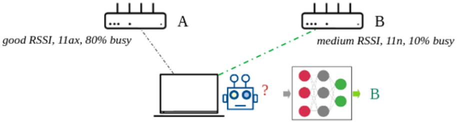
Numerical Predictions (Regression)

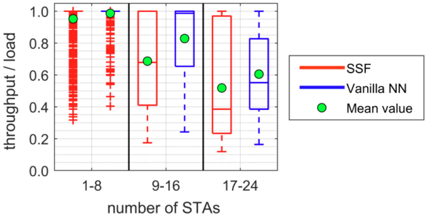
y = temperature
y = number of visitors
y = number of kW
y = price
y = number of hours

Supervised learning

- **Goal:** Training data includes x and y, so the goal is to learn $f(x) = y$.
- **Data:** We need a consolidated dataset to train the model, retraining as the dataset updates.
- **Techniques:** Time series analysis, regressions, vector machines, decision trees, artificial neural networks, etc.
- **Models:** linear regression, support vector machine (SVM), feed-forward neural networks (FFNN), convolutional neural networks (CNN), graph neural networks (GNNs), transformer, etc.

Wi-Fi example: AP selection



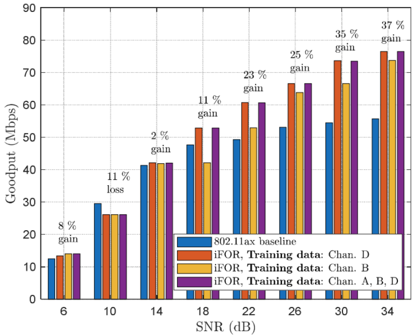


The STA uses a trained model that predicts which is the best AP to associate. The input of the model should consist of the info gathered by the STA for each AP.

Unsupervised learning

- **Goal:** Find patterns/trends/similarities in a dataset when only the features are available
- **Data:** We need a consolidated dataset to train the model, retraining it as the dataset updates
- **Models:** K-means, Self-Organizing Maps (SOMs), hidden Markov model (HMM), Auto encoders (AEs), Principal Component Analysis (PCA), Restricted Boltzmann machine (RBM), fuzzy C-means, etc.

Wi-Fi example: CSI overhead reduction

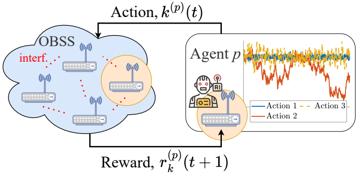


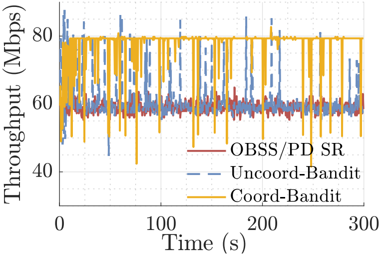
[Source] Deshmukh, Mrugen, Mahmoud Kamel, Zinan Lin, Rui Yang, Hanqing Lou, and Ismail Güvenç. "Intelligent Feedback Overhead Reduction (iFOR) in Wi-Fi 7 and Beyond." In 2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring), pp. 1-5. IEEE, 2022.

Reinforcement learning

- **Goal:** Learn from experience (relationship between actions and rewards).
- **Data:** sequence of signals (rewards) acquired by performing actions.
- **Models/algorithms:** Multi-armed Bandits, Deep Reinforcement Learning (DRL), Q-learning.

Wi-Fi example: SR optimization



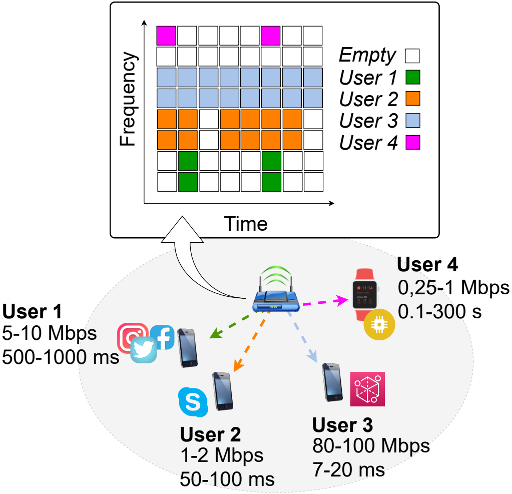


Outline

What is Resource Allocation?

Resource Allocation

- In Wi-Fi, resources primarily refer to **frequency**, **time**, and **space**, which lead to **available airtime**.
- Resource allocation is the process of distributing (or *gaining*) these resources among devices.
- Optimization goals include **network performance**, **fairness**, **efficiency**, and **quality of service (QoS)**.



Why do we need AI/ML for Wi-Fi resource allocation?

Underlying uncertainty

- Wi-Fi operates on a shared wireless medium (collisions and unbounded delays).
- Wireless communications are highly varying (high mobility, channel effects, etc.).

Increasing complexity

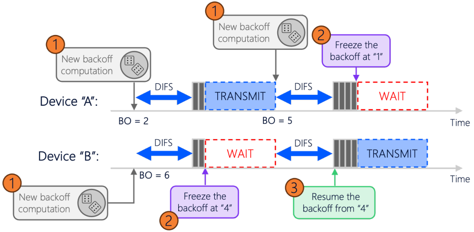
- Plethora of features and degrees of freedom (from basic CSMA/CA to advanced mechanisms MU-MIMO, or OFDMA).
- A plethora of different services and requirements makes it difficult to provide closed-form solutions.

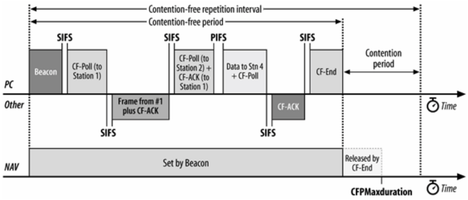
Fairness and backwards compatibility

- All devices get a reasonable share of the resources.
- Wi-Fi is backwards-compatible with older generations that might not implement advanced features.

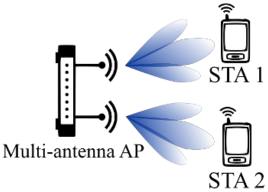
How Resources are Gained: Channel Access

- **Distributed Coordination Function (DCF):** The basic channel access protocol in Wi-Fi (CSMA/CA + random backoff).
- **Point Coordination Function (PCF):** An optional access method where the AP acts as a central controller to grant transmission opportunities.
- **Hybrid Coordination Function (HCF):** Introduced in IEEE 802.11e to enhance QoS support. Two channel access methods:
  - Enhanced Distributed Channel Access (EDCA)
  - HCF Controlled Channel Access (HCCA)

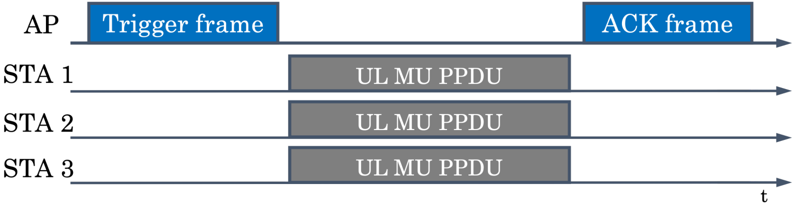




Efficient Ways of Using Resources (I)

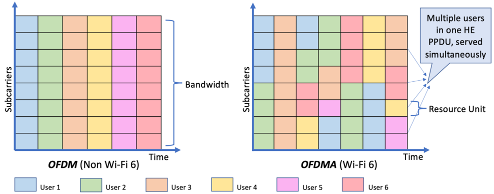


- Multi-User MIMO (MU-MIMO)
  - Allows the AP to transmit to and receive from multiple STAs simultaneously using multiple antennas.
  - Key concepts: spatial diversity (exploit different spatial paths for the signals) and precoding (combine signals constructively for the intended data stream and destructively for the data streams intended for other clients).
  - Resource allocation involves pairing STAs and allocating spatial streams.



Efficient Ways of Using Resources (II)

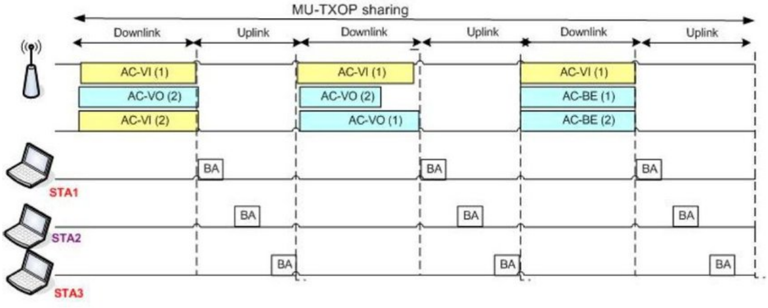
- Orthogonal Frequency-Division Multiple Access (OFDMA)
  - Divides the available bandwidth into smaller sub-channels called Resource Units (RUs), which the AP allocates to STAs.
  - Resource allocation involves determining the optimal size of RUs and assigning them to STAs based on their needs.



RU Type	20 MHz Channel	40 MHz Channel	80 MHz Channel	160 MHz Channel	80+80 MHz Channel
26-tone	9	18	37	74	74
52-tone	4	8	16	32	32
106-tone	2	4	8	16	16
242-tone	1	2	4	8	8
484-tone	-	1	2	4	4
996-tone	-	-	1	2	2
2x996-tone	-	-	-	1	1

Efficient Ways of Using Resources (III)

- Transmit Opportunity (TXOP) Sharing
  - Allows a device that has gained access to the medium (TXOP winner) to share a portion of its TXOP with other devices.
  - Resource allocation involves determining the STAs with which the TXOP is shared.



Outline

The Channel Selection Problem

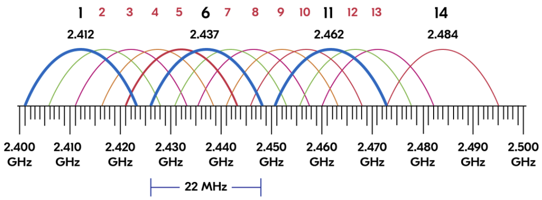
- The AP has to choose one channel:

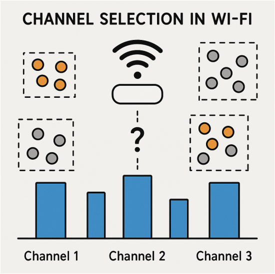
$$i \in \{1, 2, ..., N\}$$

- The reward is a value between 0 and 1, and represents the fraction of time the channel has been used over a certain interval Δ (i.e., channel occupancy, ρ):

$$r = 1 - \frac{\rho(c)}{\Delta}$$

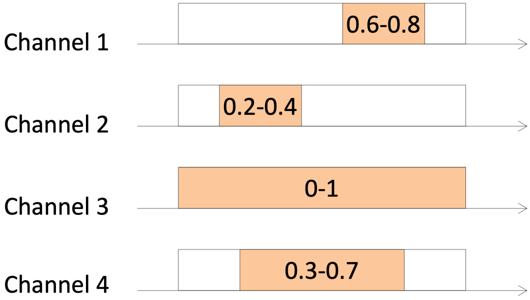
- The AP only knows the occupancy of the channel it is currently using, but cannot see the occupancy of the other channels (bandit feedback).





The Channel Selection Problem (Example)

- We consider an example where the reward distributions behind each channel can be modeled as a stochastic process.
- We try four different algorithms:
  - Pure exploration (or brute force, `BE`): Explore at the beginning and exploit later.
  - ε-greedy (`EG`): Decide to explore or exploit based on a fixed probability (ϵ).
  - UCB (`UCB`): Define bounds for rewards and explore optimistically.
  - Thompson sampling (`TS`): Build distributions from observations and sample from them.

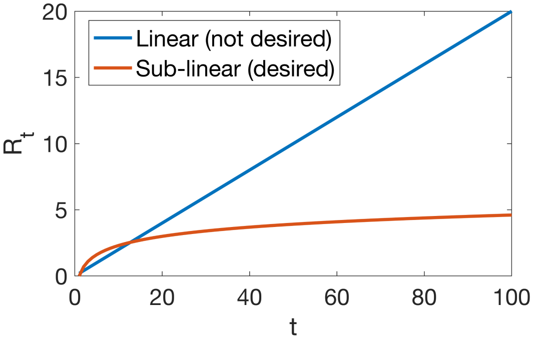


Example

`https://github.com/fwilhelmi/ml4net_course/blob/main/Examples/AAX_ChannelSelection.ipynb`

ε-greedy: Expected regret (I)

- Each channel i has an associated stochastic reward distribution.
  - μ_i is the true expected reward of channel i.
  - $\mu^* = \max_i \mu_i$ is the expected reward of the optimal channel i^*.
  - $\Delta_i = \mu^* - \mu_i \geq 0$ is the *sub-optimality gap* for channel i.
- We want to analyze the regret of the ϵ-greedy algorithm.



ε-greedy: Expected regret (II)

At each time step t:

- **With probability** $1-\epsilon$, the empirically best channel is selected

$$i(t) = i_{best}(t-1) = \arg\max_i \hat{\mu}_i(t-1), \tag{1}$$

  where $\hat{\mu}_i(t-1)$ is the average reward of channel i up to time $t-1$.
- **With probability** ϵ, a random channel is selected

$$i(t) \sim \mathcal{U}(\mathcal{N}) \tag{2}$$

Therefore, the expected reward at time t, $\mathbb{E}[r_t]$, is:

$$\begin{aligned}\mathbb{E}[r_t] &= (1-\epsilon)\mathbb{E}[r_t|\text{exploitation}] + \epsilon\mathbb{E}[r_t|\text{exploration}] \\ &= (1-\epsilon)\mathbb{E}[\mu_{i_{best}(t-1)}] + \epsilon\left(\frac{1}{N}\sum_{i=1}^{N}\mu_i\right)\end{aligned}$$

ε-greedy: Expected regret (III)

The expected regret R_t at time t is:

$$
\begin{aligned}
\mathbb{E}[R_t] &= \mu^* - \mathbb{E}[r_t] \\
&= \mu^* - (1-\epsilon)\mathbb{E}[\mu_{i_{best}(t-1)}] - \frac{\epsilon}{|\mathcal{C}|}\sum_{i=1}^{|\mathcal{C}|}\mu_i \\
&= (1-\epsilon)(\mu^* - \mathbb{E}[\mu_{i_{best}(t-1)}]) + \epsilon\left(\mu^* - \frac{1}{N}\sum_{i=1}^{N}\mu_i\right) \\
&= (1-\epsilon)\mathbb{E}[\Delta_{i_{best}(t-1)}] + \frac{\epsilon}{N}\sum_{i=1}^{N}\Delta_i
\end{aligned}
$$

where $\Delta_i = \mu^* - \mu_i$.

ε-greedy: Expected regret (IV)

The expected cumulative regret up to time T is:

$$\begin{aligned}
\mathbb{E}[R(T)] &= \sum_{t=1}^{T} \mathbb{E}[\mu^* - r_t] \\
&= \sum_{t=1}^{T} \left[(1-\epsilon)\mathbb{E}[\Delta_{i_{best}(t-1)}] + \frac{\epsilon}{N} \sum_{i=1}^{N} \Delta_i \right] \\
&= (1-\epsilon) \sum_{t=1}^{T} \mathbb{E}[\Delta_{i_{best}(t-1)}] + T \frac{\epsilon}{N} \sum_{i=1}^{N} \Delta_i
\end{aligned}$$

- The first term, $(1-\epsilon)\sum_{t=1}^{T} \mathbb{E}[\Delta_{i_{best}(t-1)}]$, is the regret during exploitation of potentially sub-optimal arms and grows sub-linearly with T.
- The second term, $T\frac{\epsilon}{N}\sum_{i=1}^{N} \Delta_i$, is the regret of exploring sub-optimal actions and grows linearly with T.

ε-greedy: Expected regret (V)

Therefore, for a constant $\epsilon > 0$, the expected cumulative regret is:

$$\mathbb{E}[R(T)] \leq O\left(\sum_{i:\mu_i<\mu^*} \frac{\log T}{\Delta_i}\right) + O(\epsilon T)$$

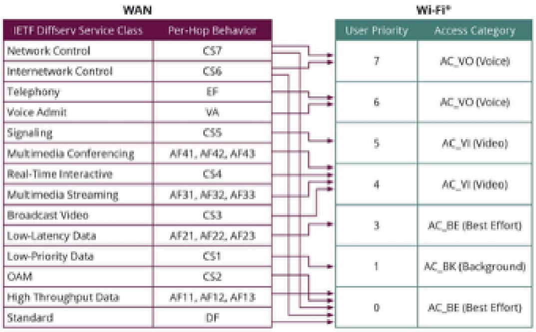
To achieve sub-linear regret, ϵ needs to decrease over time, e.g., $\epsilon_t = \epsilon_0/t$ or $\epsilon_t = \epsilon_0/\sqrt{t}$, so that we can focus more on the best action:

$$\mathbb{E}[R(T)] \leq O\left(\sum_{i:\mu_i<\mu^*} \frac{\log T}{\Delta_i}\right)$$

Outline

Enhanced Distributed Channel Access (EDCA)

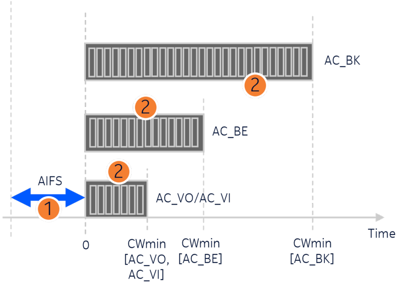
- Wi-Fi provides Quality of Service (QoS) through the Enhanced Distributed Channel Access (EDCA), an extension of the DCF that was defined in the IEEE 802.11e specification.
- EDCA enforces priorities to different traffic types or Access Categories (AC) through the channel access:
  - Voice (AC_VO)
  - Video (AC_VI)
  - Best effort (AC_BE)
  - Background (AC_BK)



RFC 8325 DSCP to Wi-Fi ACs mapping [Source: Wi-Fi CERTIFIED QoS Management]

Contention Window (CW)

- The Contention Window (CW) is used to calculate the random backoff used for accessing the channel.
- The backoff is randomly chosen between 0 and CW.
- An exponential phase is entered if packet losses are experienced:
  - When a loss occurs, the CW is doubled (up to CWmax).
  - When a successful transmission is done, the CW is restarted to CWmin.



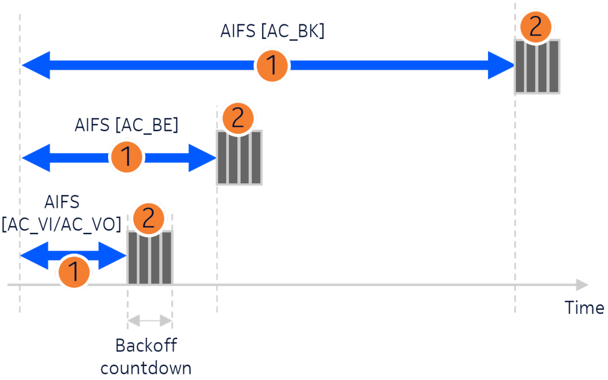
Example

CWmin [AC_VI/AC_VO] = 7, CWmin [AC_BE] = 31, and CWmin [AC_BK] = 63 leads to expected random backoffs of 3.5, 15.5, and 31.5 for each considered AC, respectively.[a]

[a] The expected backoff is computed as CW/2.

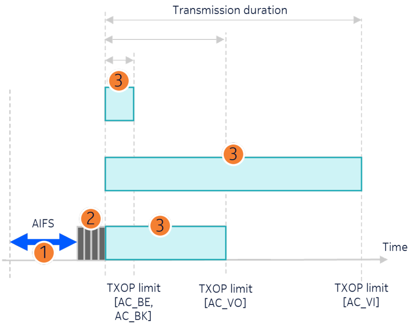
Arbitrary Inter-Frame Space (AIFS)

- The AIFS is a time interval to be respected before starting the channel access procedure that precedes the transmission of a new frame.
- Configuring a small AIFSN value allows a device to start decreasing its random backoff counter before other AC, thus shortening its total waiting time.



Maximum TXOP duration

- The TXOP limit defines the maximum duration that a device can occupy the channel for transmitting.
- The TXOP limit has a strong implication on the airtime achieved by each traffic type (airtime = time spent occupying the channel).



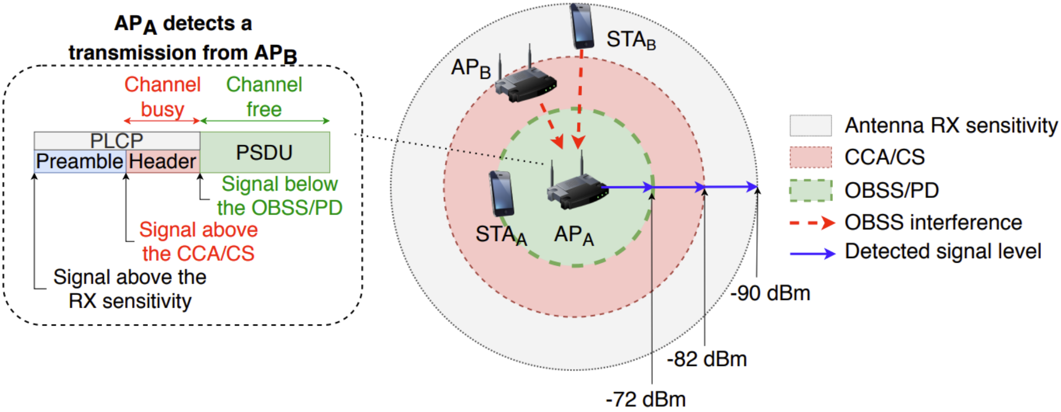
Example

For CWmin [AC_VI/AC_VO] = 7, CWmin [AC_BE] = 31, and CWmin [AC_BK] = 63, assigning the TXOP limits TXOP[AC_VI] = 3 ms TXOP[AC_VO] = 2 ms, TXOP[AC_BE, AC_BK] = 1 ms would lead each AC enjoying the following airtime values:

- Video (AC_VI): 57%, Voice (AC_VO): 38%, Best effort (AC_BE): 3.8%, Background (AC_BK): 1.2%

Outline

Spatial reuse in Wi-Fi



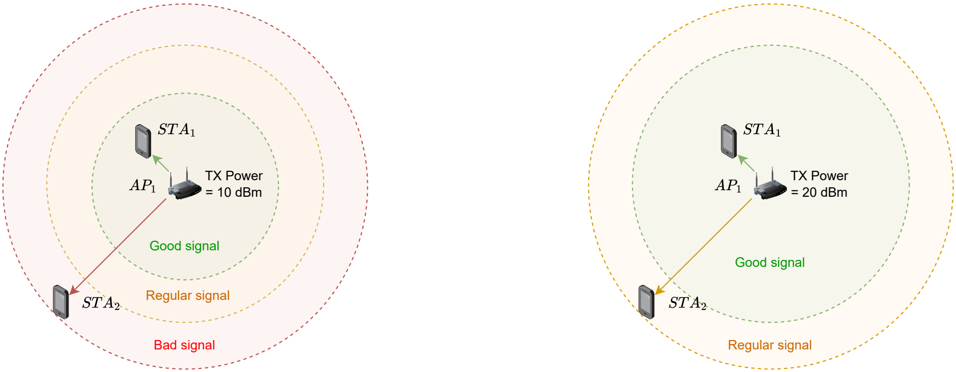
- CCA (Clear Channel Assessment): First SD threshold check.
- OBSS/PD: Second (more relaxed) SD threshold check.

Transmission power (I)

- The transmission power employed by a given transmitter determines the strength perceived at the receiver
  - A better SINR allows for adopting higher-order MCS (= more capacity)
- Transmission power in wireless communication is often measured in:
  - Watts (W) or milliwatts (mW)
  - dBm (decibels relative to one milliwatt)

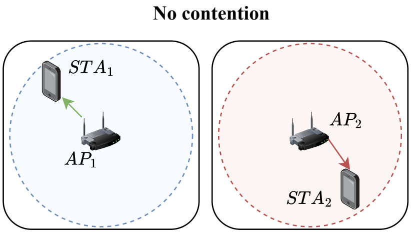
SINR [dB]	Modulation	Coding Rate	TP [b/s/Hz]
∼ -5.0	No use		0
-5.0 ∼ -1.9	QPSK	1/8	0.25
-1.9 ∼ 1.8	QPSK	1/4	0.50
1.8 ∼ 3.8	QPSK	1/2	1.00
3.8 ∼ 7.1	QPSK	2/3	1.33
7.1 ∼ 9.3	16QAM	1/2	2.00
9.3 ∼11.3	16QAM	2/3	2.67
11.3 ∼14.5	64QAM	1/2	3.00
14.5 ∼17.2	64QAM	2/3	4.00
17.2 ∼19.5	64QAM	0.81	4.86
19.5 ∼	64QAM	7/8	5.25

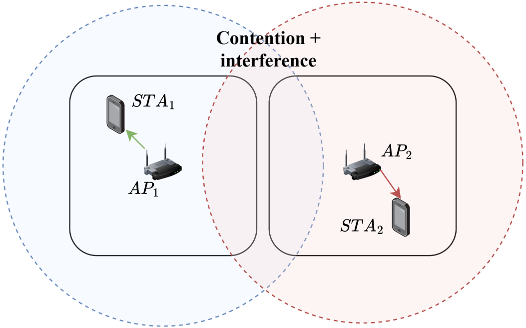
Transmission power (II)



- Higher transmission power allows reaching longer distances.
- The signal strength (thus capacity) is also benefited from higher transmission powers.

Transmission power (III)

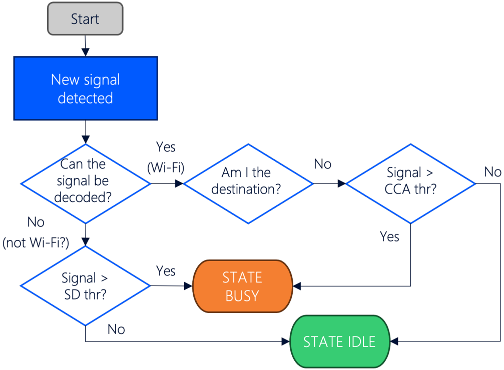




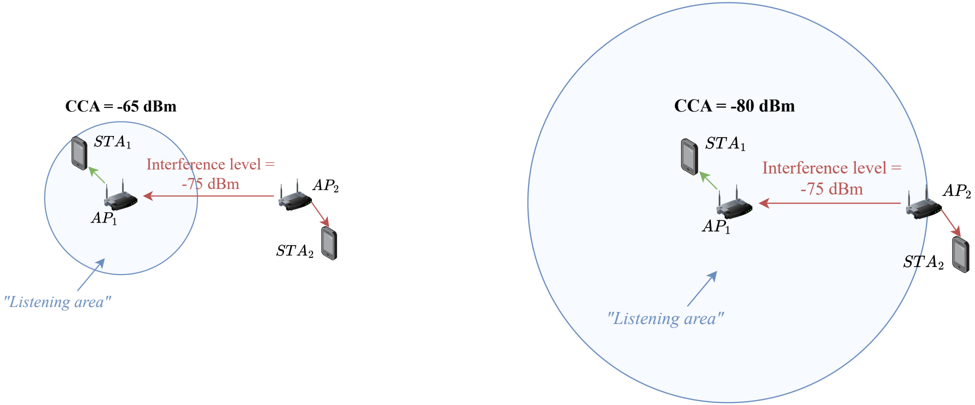
- Higher transmission power leads to higher interference, thus potentially less capacity.
- Contention also increases when the transmission power is increased.

Signal detection (I)

- As part of the CSMA/CA operation, the channel is sensed to assess whether it is idle or busy.
  - "Collision avoidance"
- Clear Channel Assessment (CCA) is employed.
  - Energy Detect (ED) threshold: Check non-Wi-Fi signals. Typically, −62 dBm.
  - Signal Detect (SD) threshold: Check Wi-Fi signals. Typically, −82 dBm.



Signal detection (II)



- Higher CCA/SD: less contention, more simultaneous transmissions, more interference, less capacity
- Lower CCA/SD: more contention, less simultaneous transmissions, less interference, more capacity

Reinforcement-Learning-based Spatial Reuse

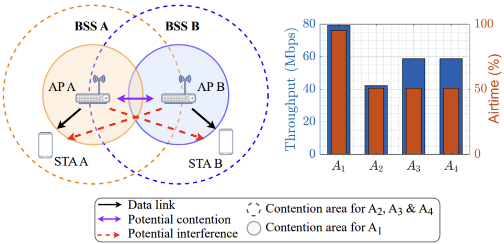
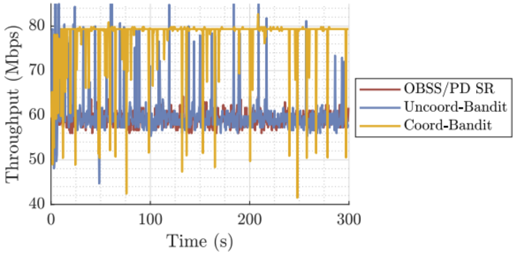


Figure 2: Considered 2-BSS toy deployment and mean performance of $A_1 = \{10, -72\}$ dBm, $A_2 = \{10, -82\}$ dBm, $A_3 = \{20, -72\}$ dBm, $A_4 = \{20, -82\}$ dBm.



Actions: Choose the best CCA/SD & transmit power values (e.g., {-82 dBm & 20 dBm}, {-72 dBm & 10 dBm}, {-62 dBm & 5 dBm})
Reward: Performance goal (throughput, delay, **airtime**, **number of TXOPs**, etc.)

Wilhelmi, F., Cano, C., Neu, G., Bellalta, B., Jonsson, A., & Barrachina-Muñoz, S. (2019). Collaborative spatial reuse in wireless networks via selfish multi-armed bandits. Ad Hoc Networks, 88, 129-141.