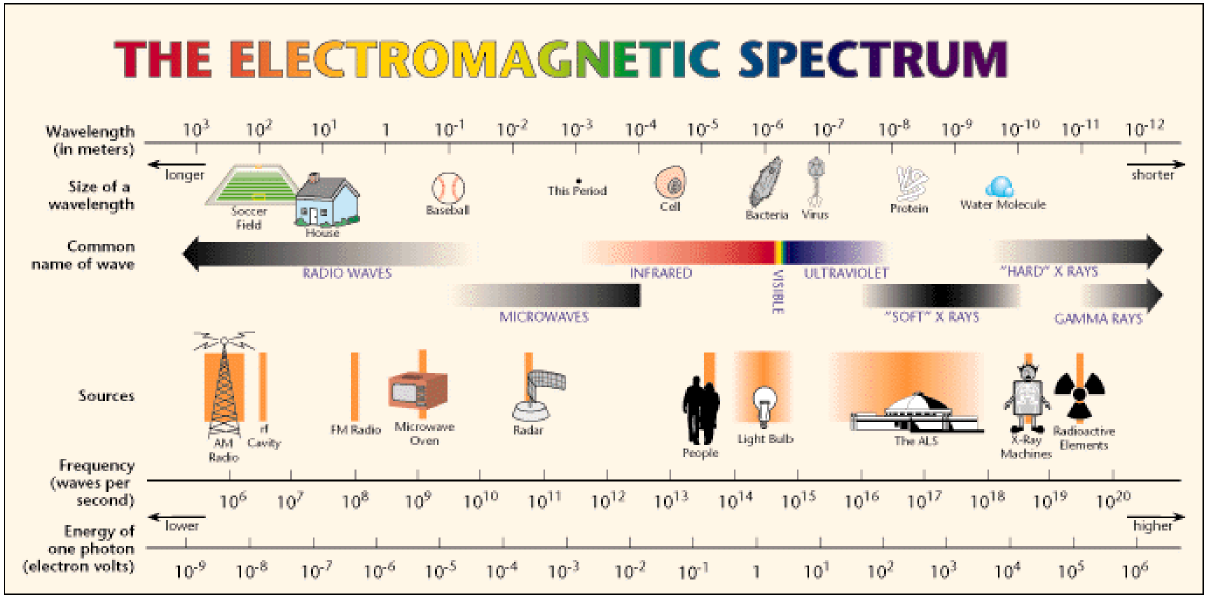
Cognitive Radio for Dynamic Spectrum Access

Qing Zhao

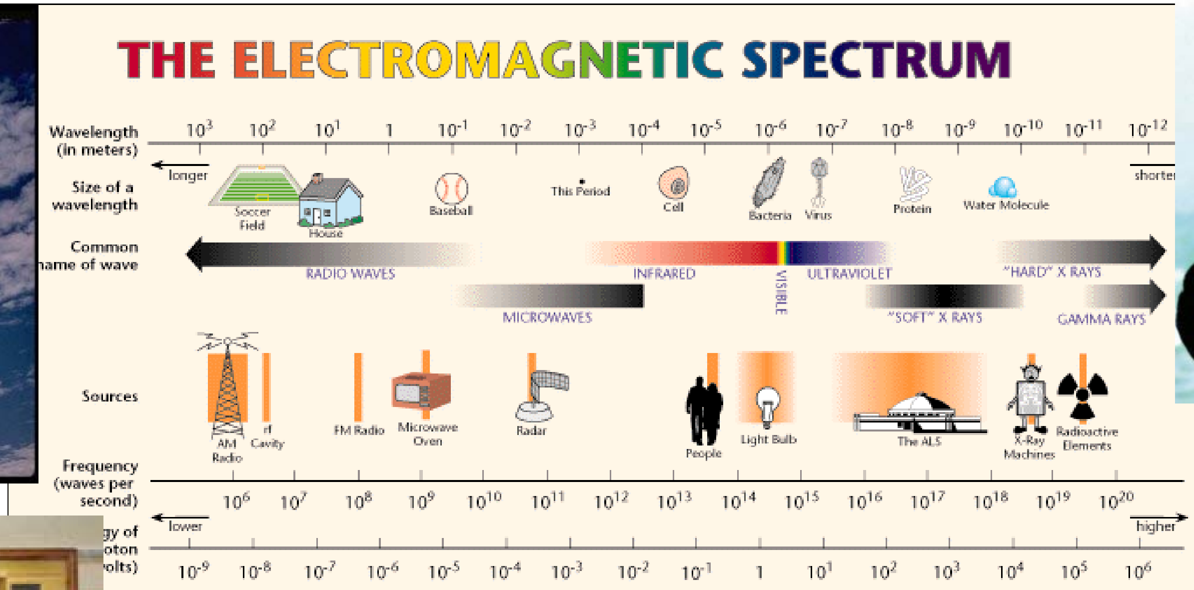
University of California at Davis

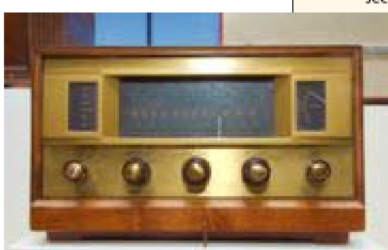
- *A Taxonomy of Dynamic Spectrum Access* *(Zhao&Sadler:07SPM)*
- *Technical Challenges in Spectrum Overlay*

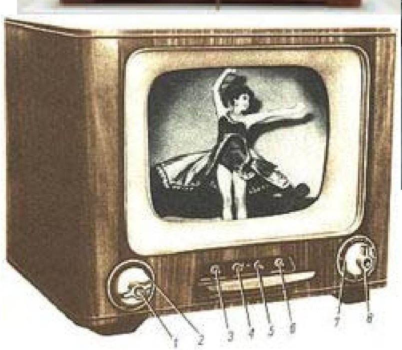
Limited Supply



Growing Demand







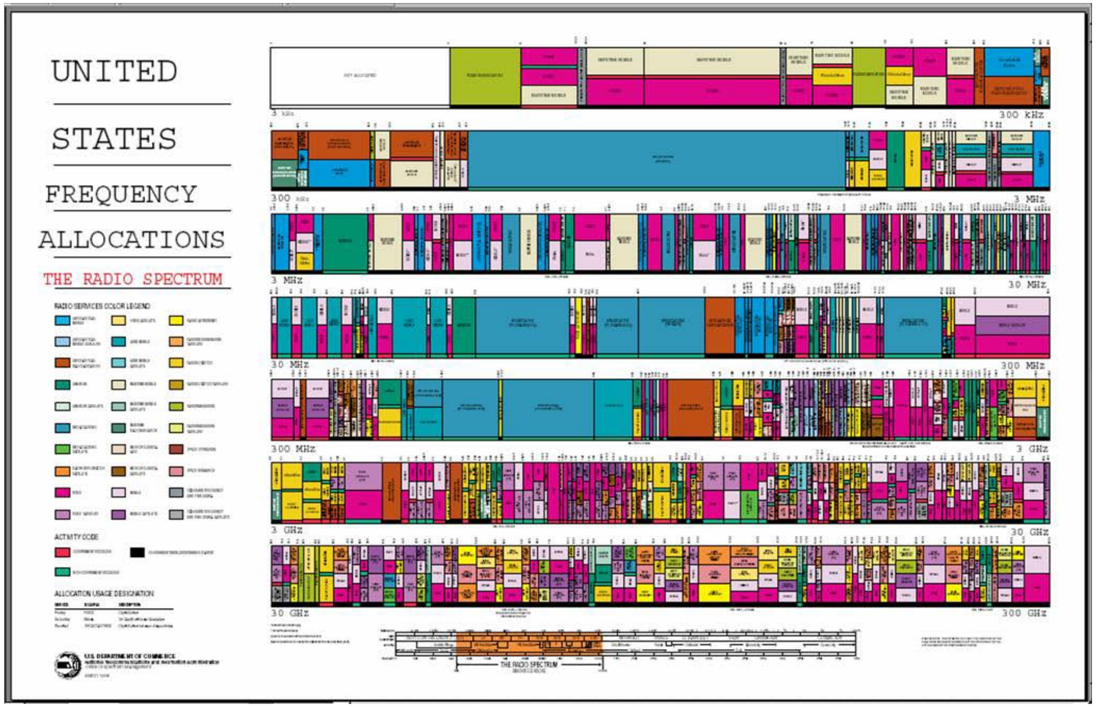
Regulation in 1912-1927: Open to All

Herbert Hoover

The Secretary of Commerce…and Under-Secretary of Everything Else!

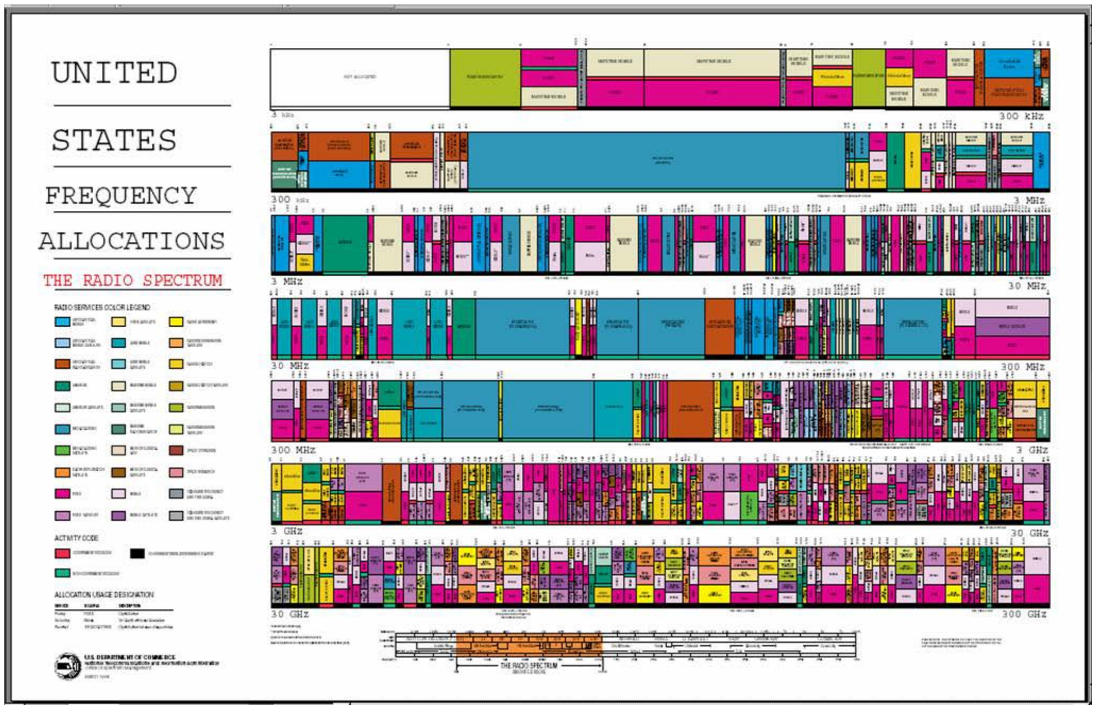
- Agency: Department of Commerce.
- Service: AM radio broadcasting.
- Limited power: cannot deny license to anyone.

Since 1927: Tight Control by FCC/FRC



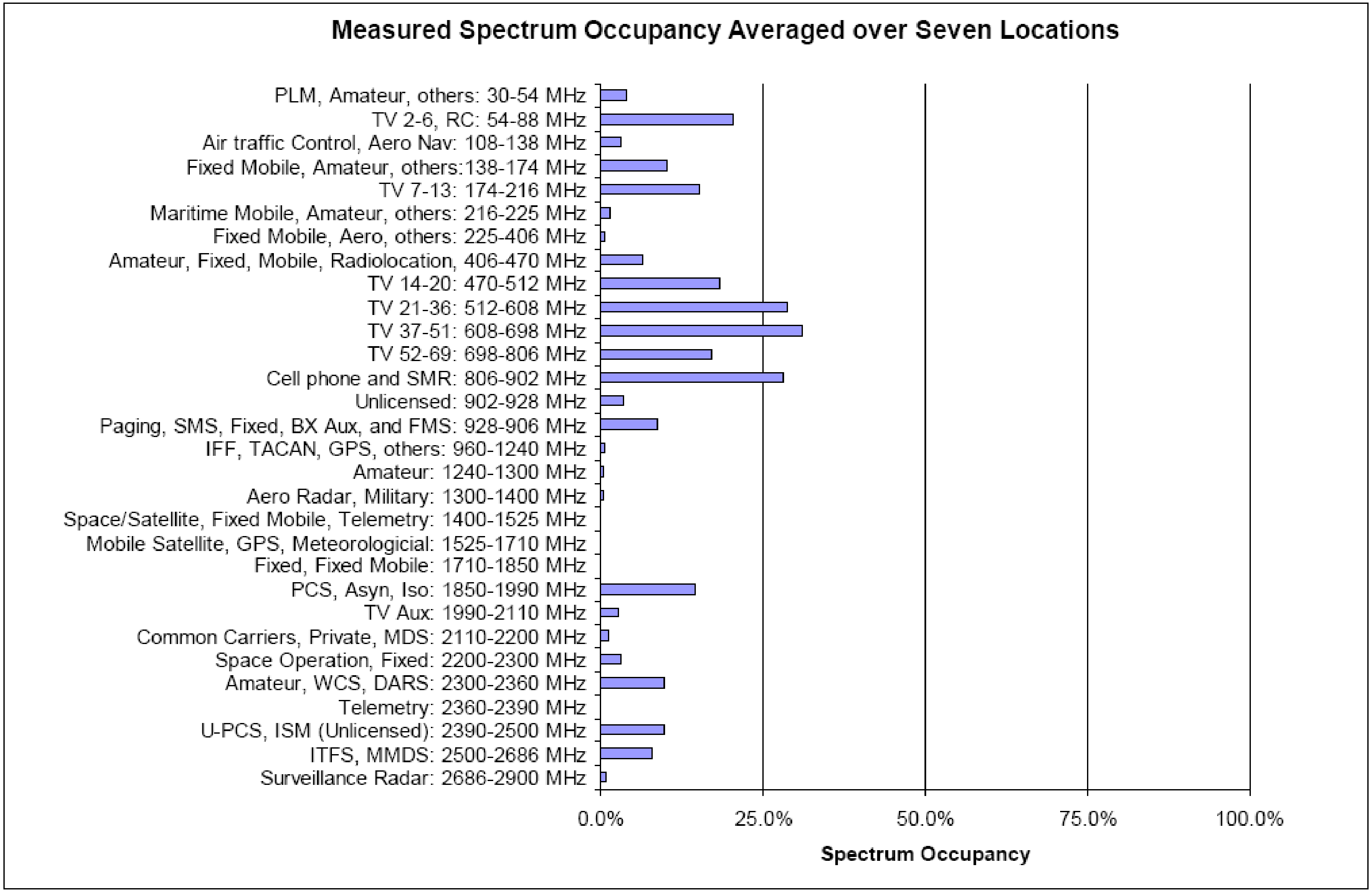
- Federal Communications Commission (FCC).
- Controls all non-Federal Government use of the spectrum.

FCC Policy & Spectrum Scarcity



- Centralized static allocation
- Little sharing
- Little flexibility

Spectrum Underutilization

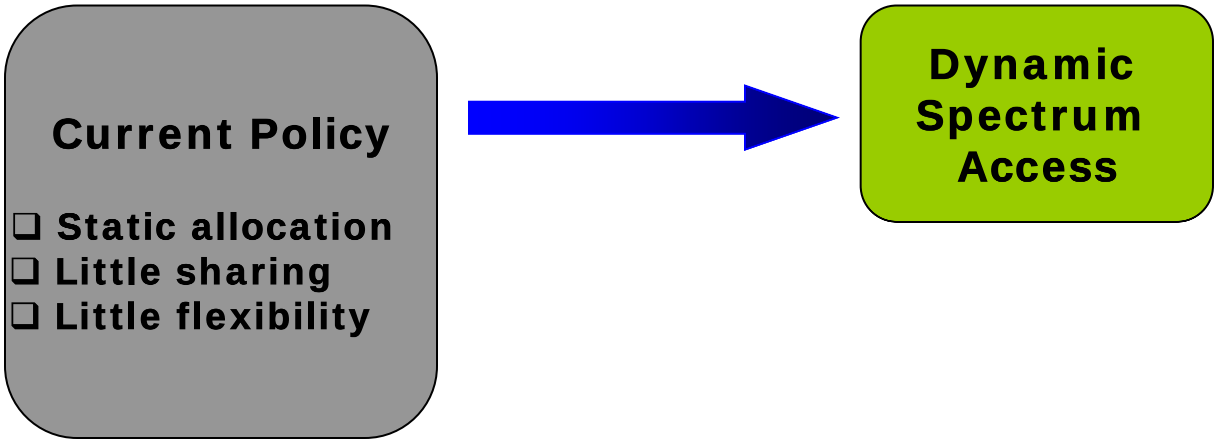


(Credit: SSC)

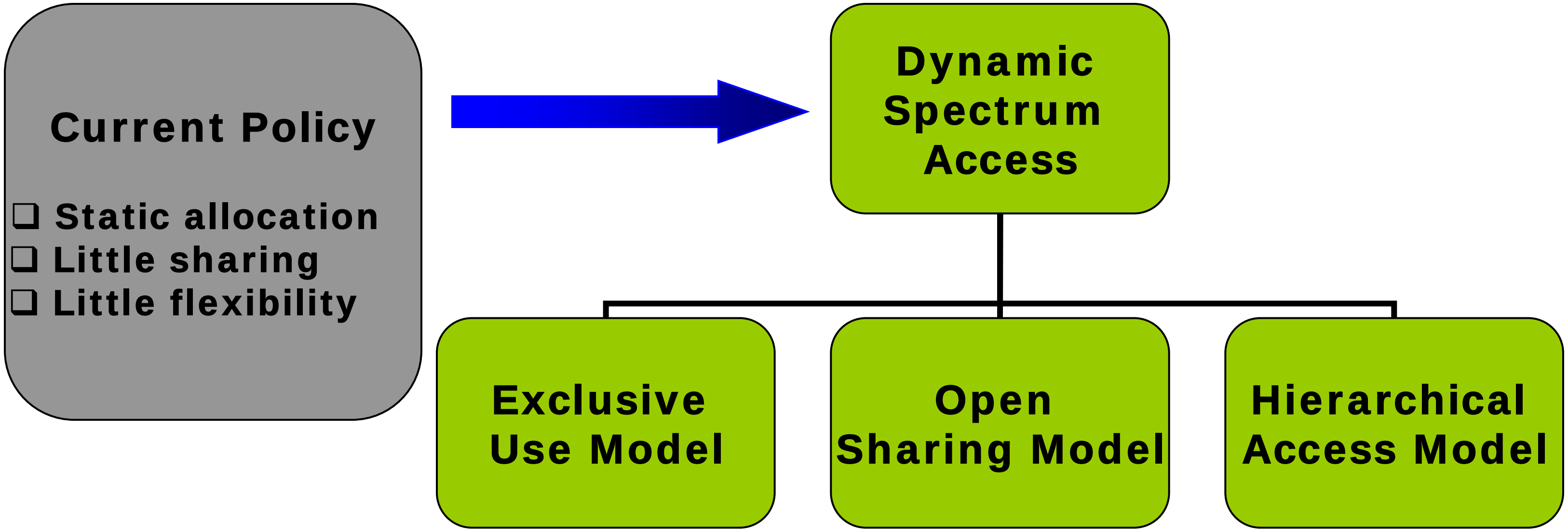
Diverse Ideas, Confusing Terms

- Dynamic spectrum access
- Dynamic spectrum allocation
- Spectrum property rights
- Spectrum commons
- Opportunistic spectrum access
- Spectrum pooling
- Spectrum underlay
- Spectrum overlay
- Cognitive radio
- ...

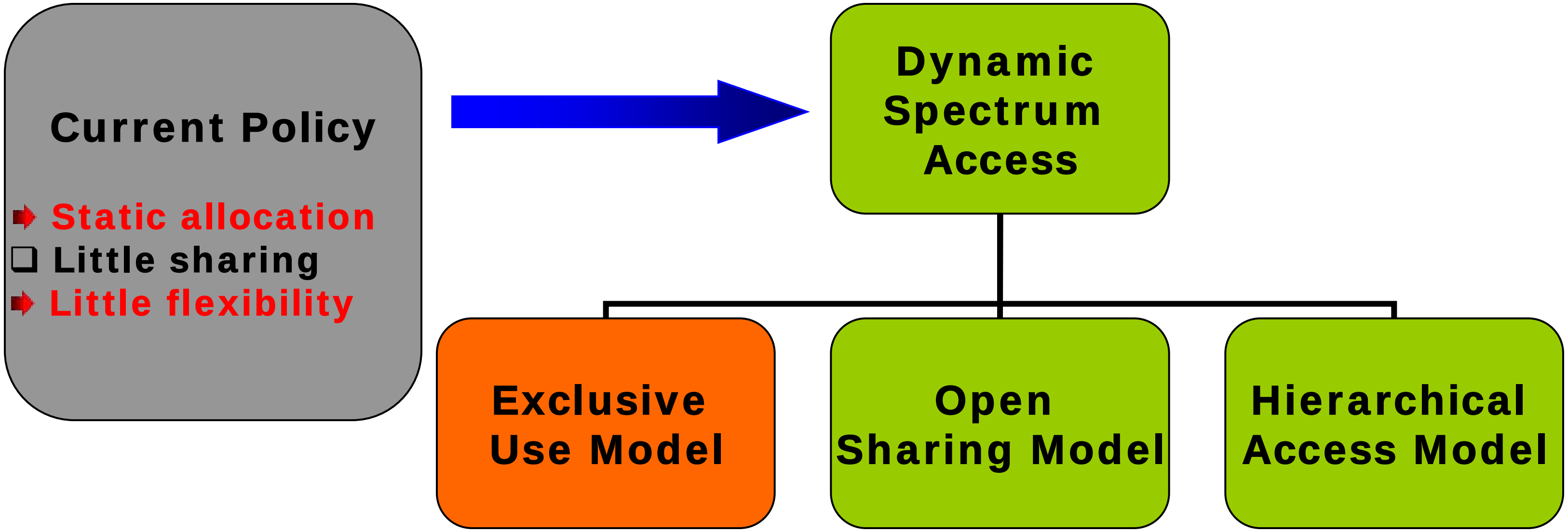
A Taxonomy of DSA



Three DSA Models

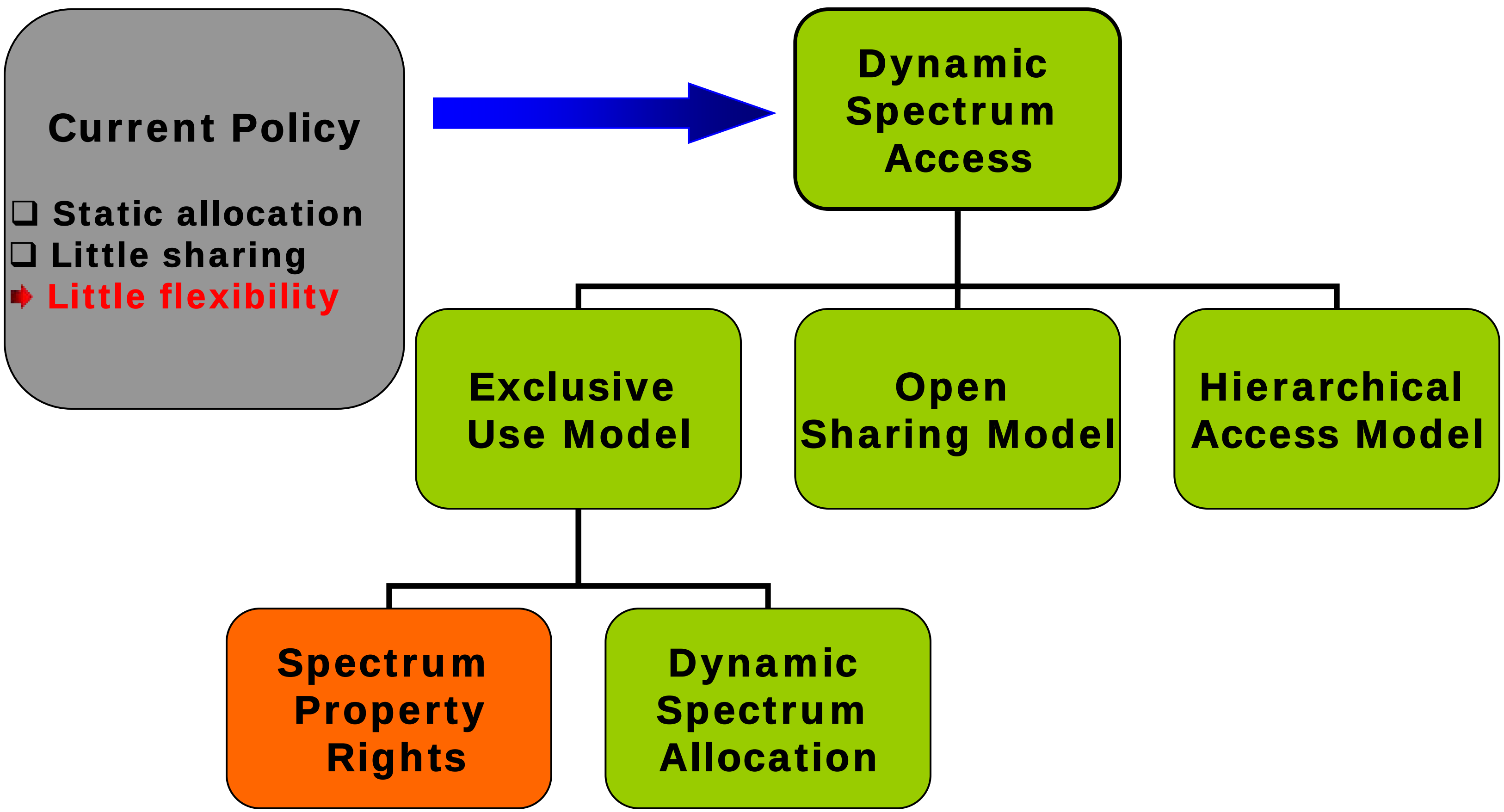


Exclusive Use Model



- Maintains the basic structure: license for exclusive use.
- Introduces flexibility in allocation and spectrum usage.

Spectrum Property Rights



- Price mechanism: allows selling and trading spectrum
- Market determines the most profitable use of spectrum

Nobel Prize Winning Idea

Ronald H. Coase

Nobel Prize Laureate in Economics (1991)

***R. Coase,* "The federal communications commission," 1959.**

Coase Theorem

Ronald H. Coase

Nobel Prize Laureate in Economics (1991)

Coase Theorem: *All government allocations of a public good are equally efficient in the absence of transaction costs.*

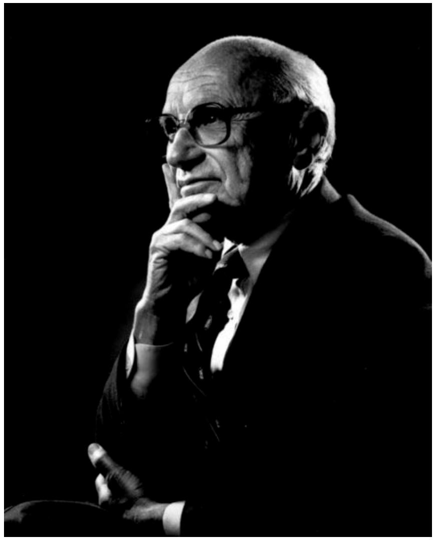
***R. Coase,* "The federal communications commission," 1959.**

Coase Theorem

Coase Theorem: *All government allocations of a public good are equally efficient in the absence of transaction costs.*

Ronald H. Coase

Nobel Prize Laureate in Economics (1991)



Milton Friedman

Nobel Prize Laureate in Economics (1976)

George J. Stigler

Nobel Prize Laureate in Economics (1982)

***R. Coase,* "The federal communications commission," 1959.**

Coase Theorem

Ronald H. Coase

Nobel Prize Laureate in Economics (1991)

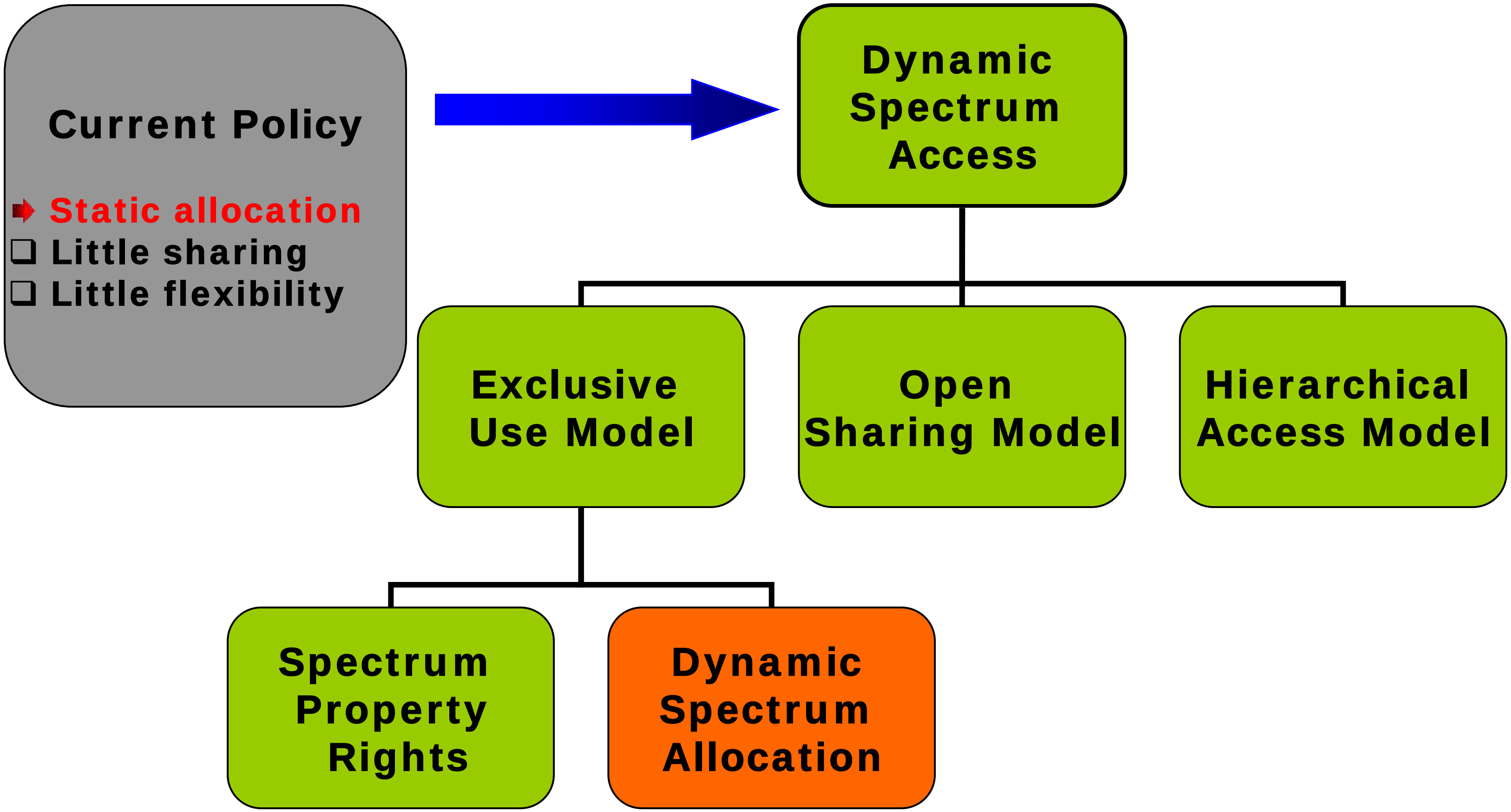
Coase Theorem: *All government allocations of a public good are equally efficient in the absence of transaction costs.*

Government Regulation: *not to find the most efficient allocation, but to minimize transaction costs.*

Spectrum Property Rights: *Allow licensees to sell and trade spectrum and freely choose technology.*

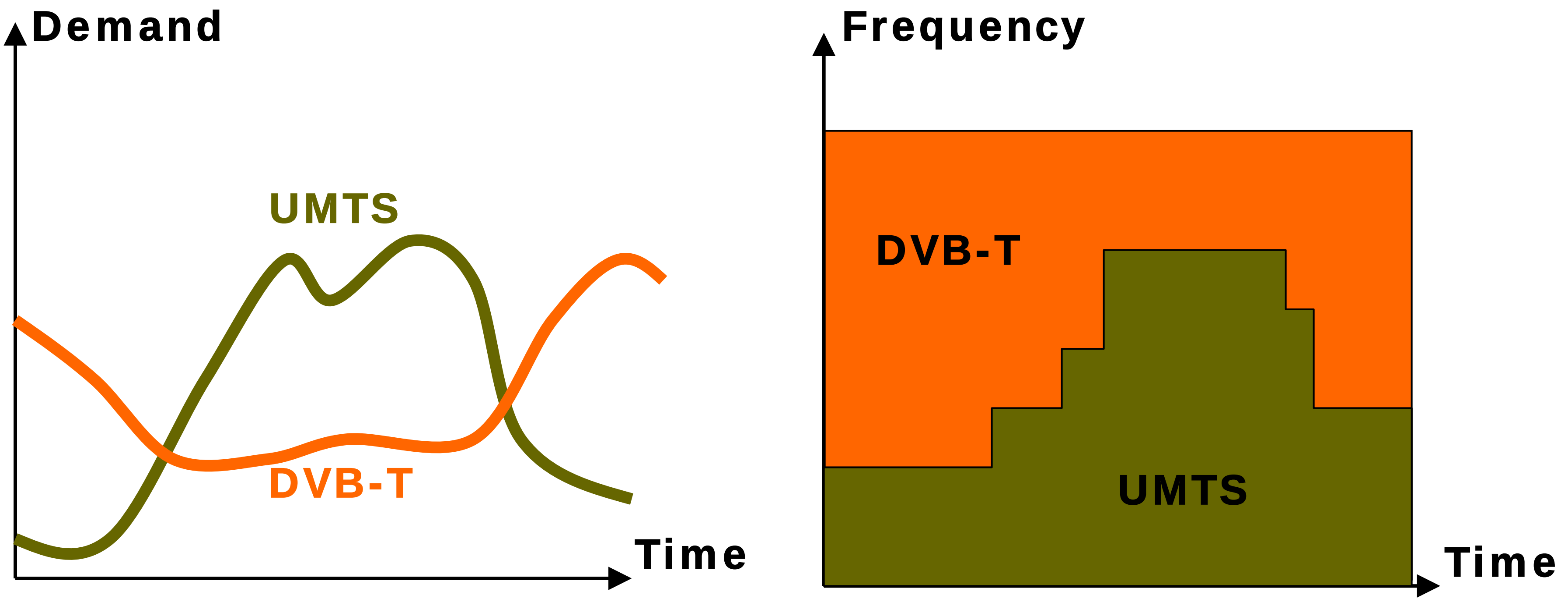
***R. Coase,* "The federal communications commission," 1959.**

Dynamic Spectrum Allocation



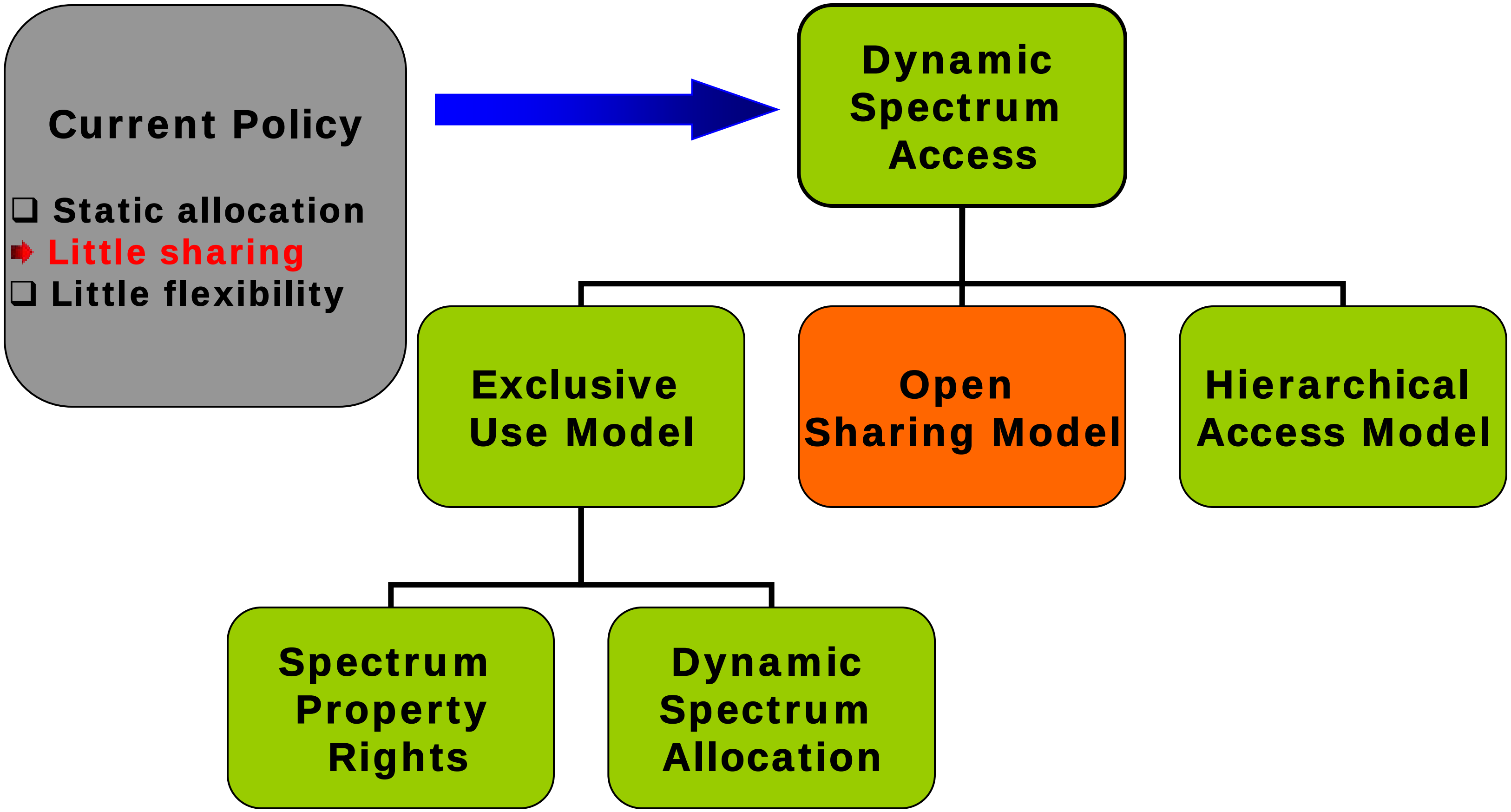
- Dynamic spectrum assignment to different services
- Exploiting spatial and temporal traffic statistics

Dynamic Spectrum Allocation



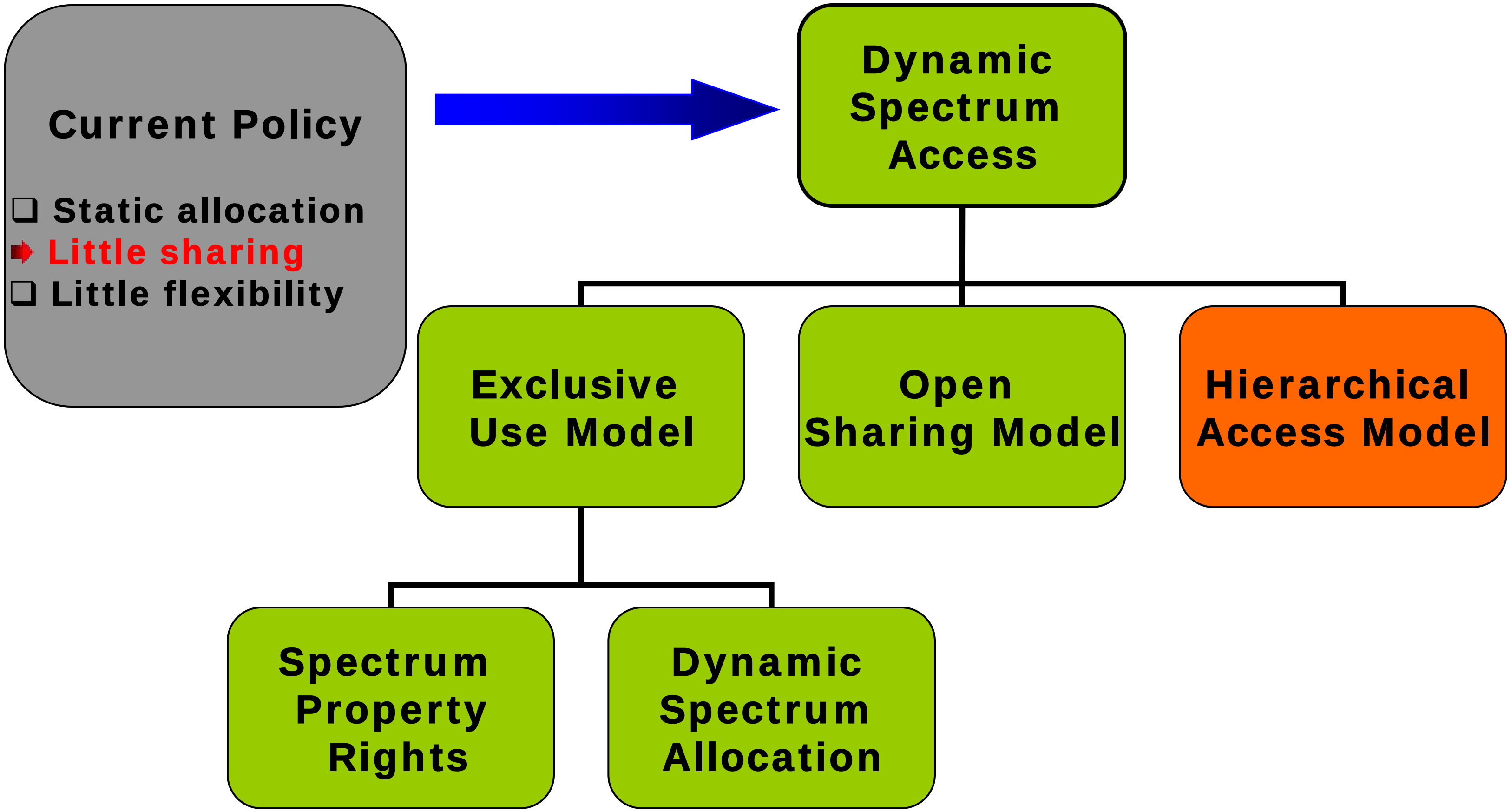
(Xu&etal:00, Leaves&etal:04)

Open Sharing Model



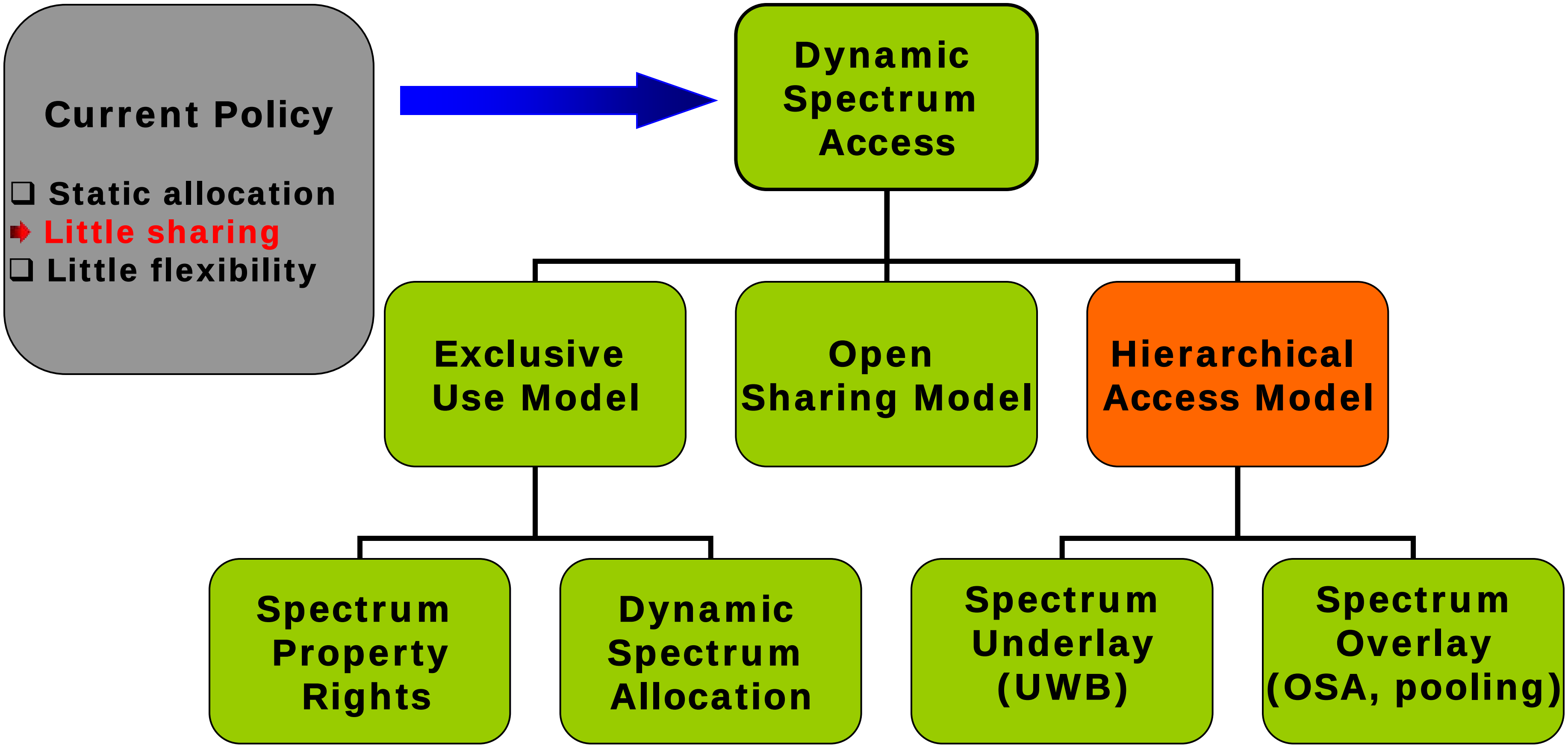
- Open sharing among peer users (spectrum commons)
- Draws support from the success of unlicensed ISM bands

Hierarchical Access Model



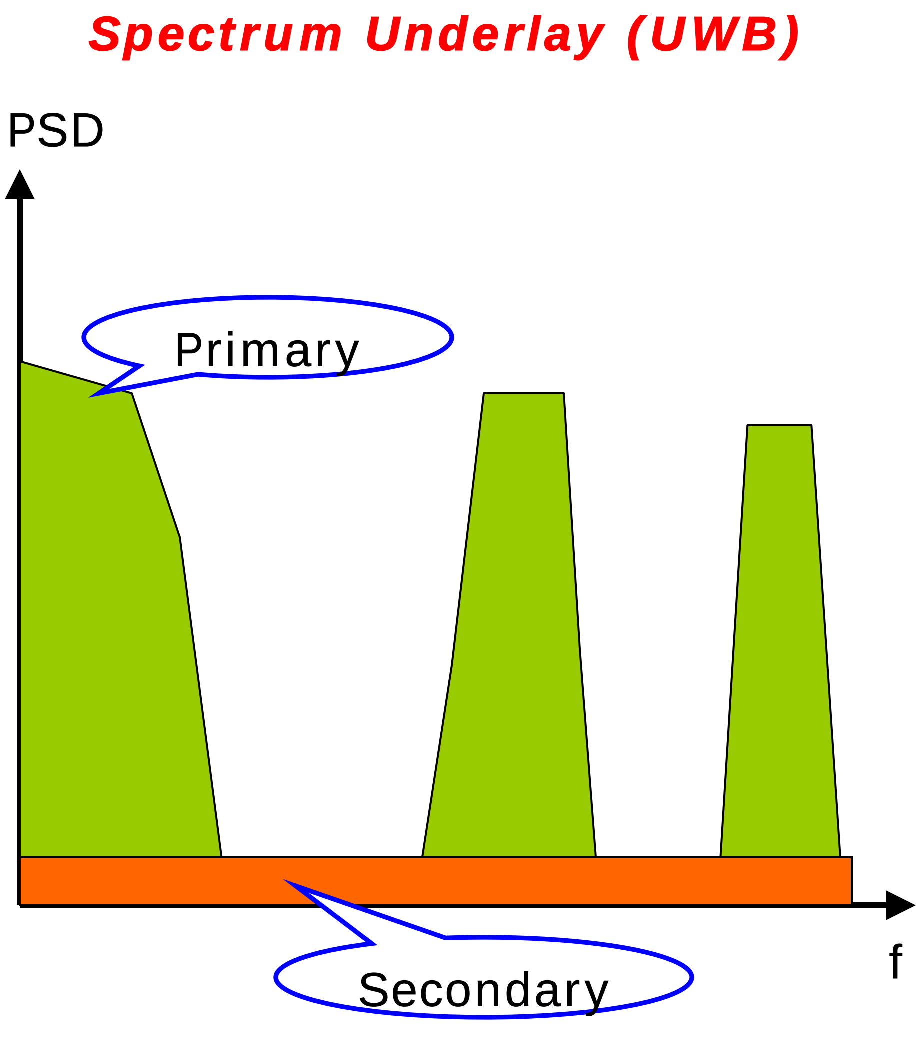
- Hierarchical access with primary and secondary users
- sharing with limited interference to primary users (licensees)

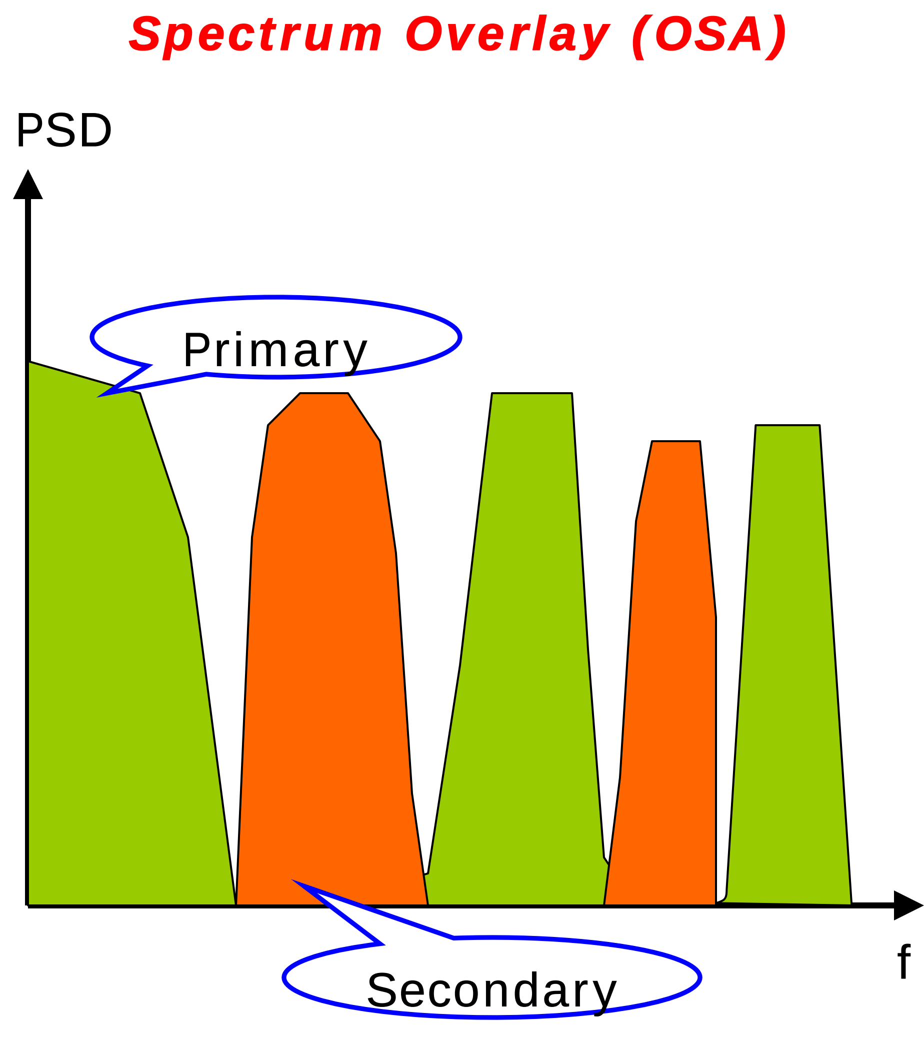
Dynamic Spectrum Access



- Spectrum underlay: constraint on transmission power
- Spectrum overlay: constraint on when and where to transmit

Underlay vs. Overlay

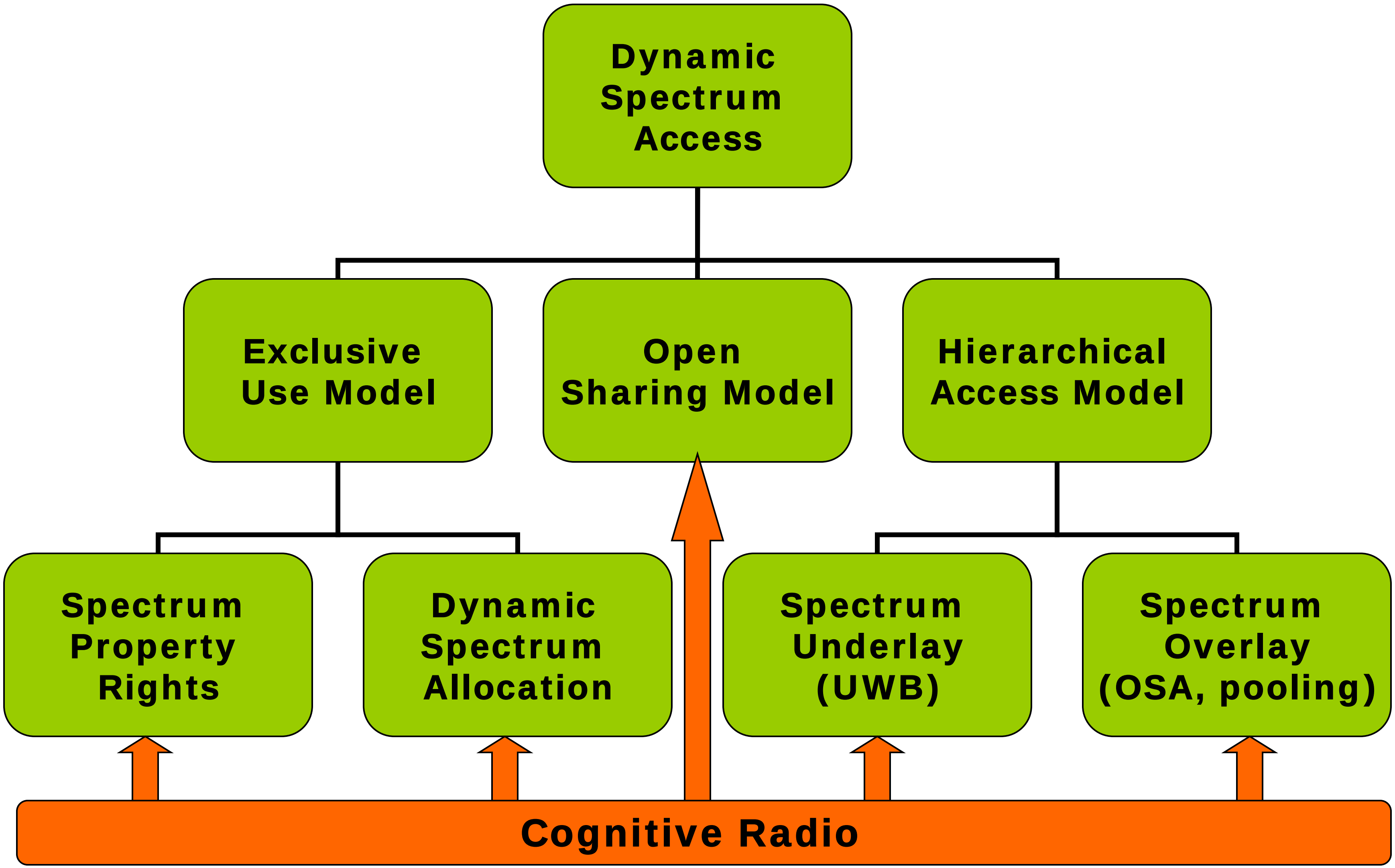




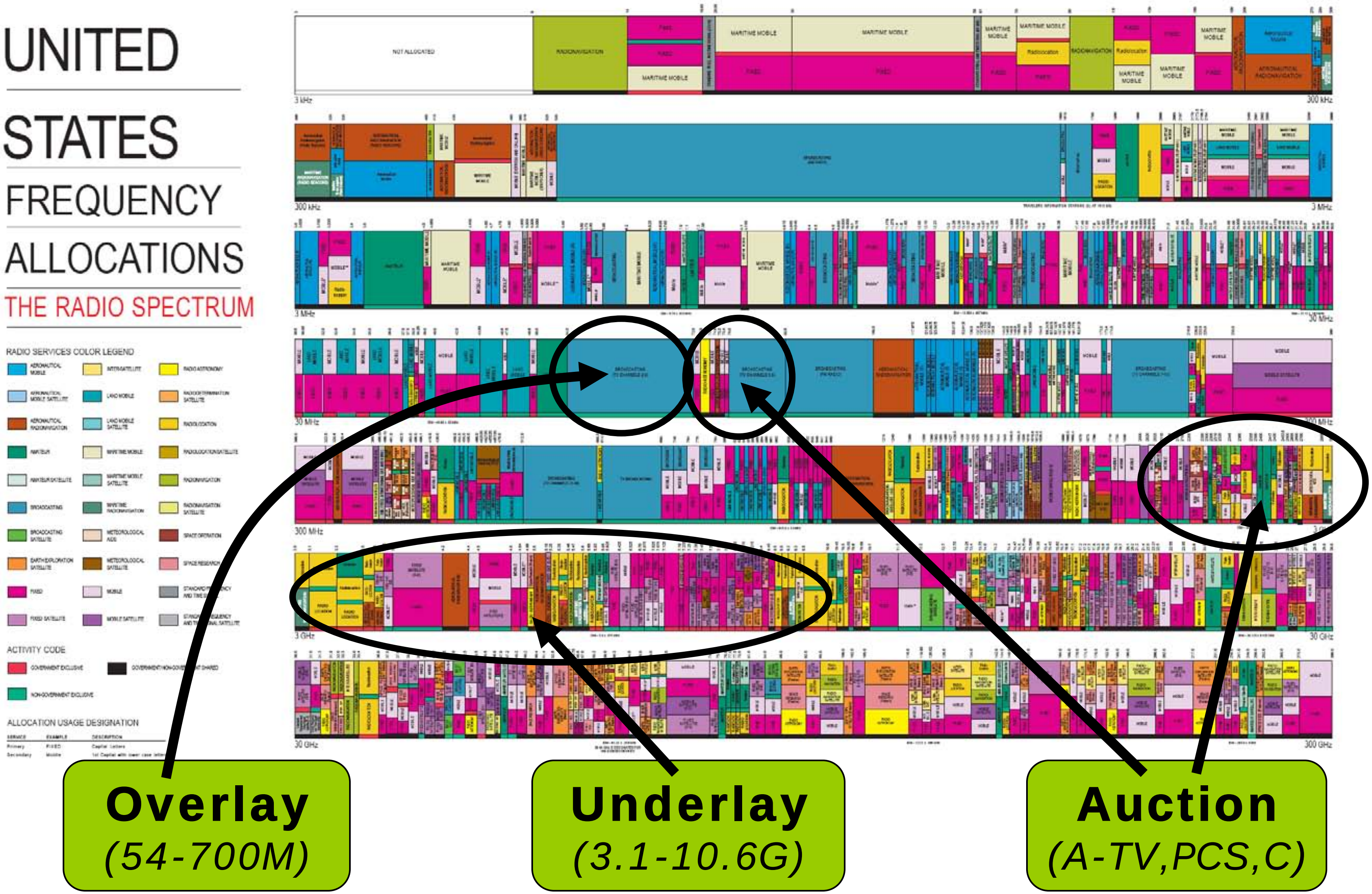
Cognitive Radio

- **Software Defined Radio**
  - Promoted by Mitola in 1991
  - A multiband radio reconfigurable through software
- **Cognitive Radio**
  - Promoted by Mitola in 1998
  - Built upon a software defined radio platform
  - Autonomously reconfigurable through learning
  - *Applications not limited to DSA*

Cognitive Radio: The Physical Platform



Towards Dynamic Spectrum Access

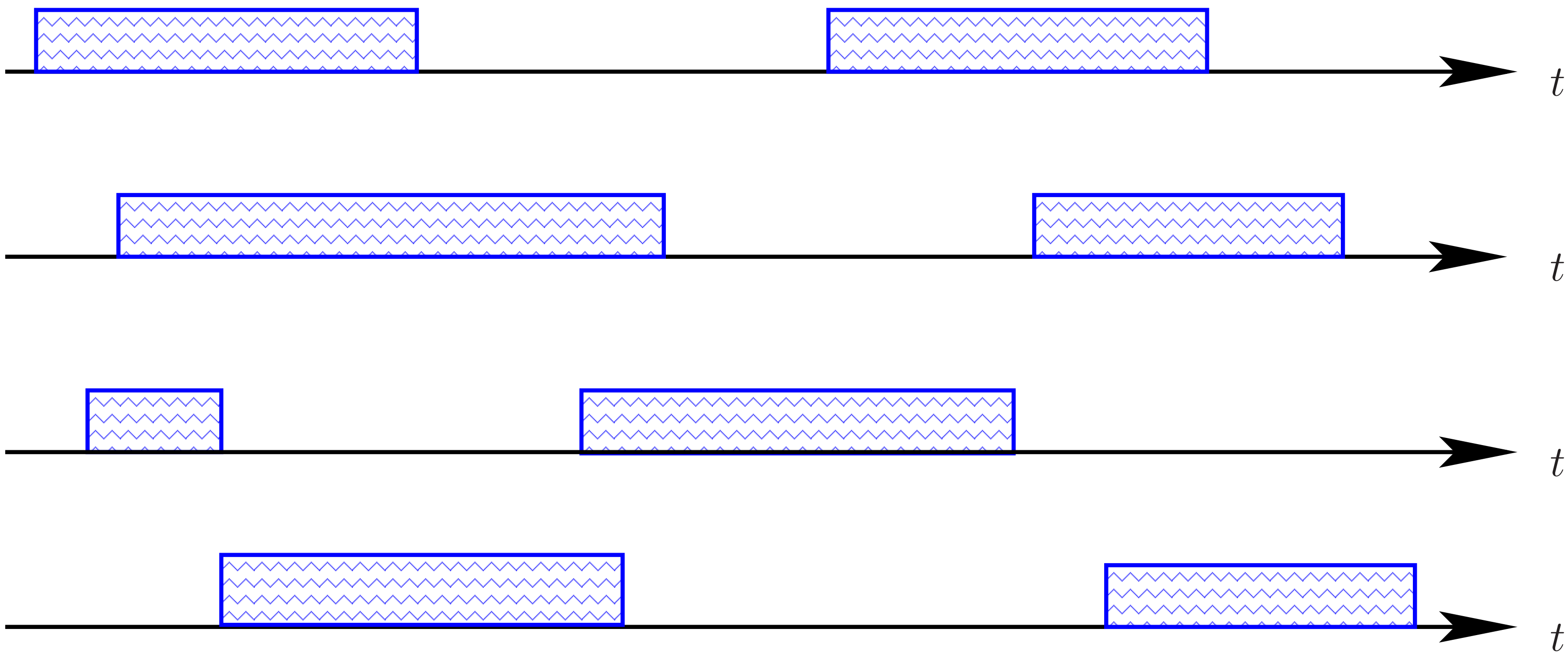


Technical Challenges in Spectrum Overlay

- *Quickest search of spectrum opportunity*
- *Distributed learning for spectrum sharing*

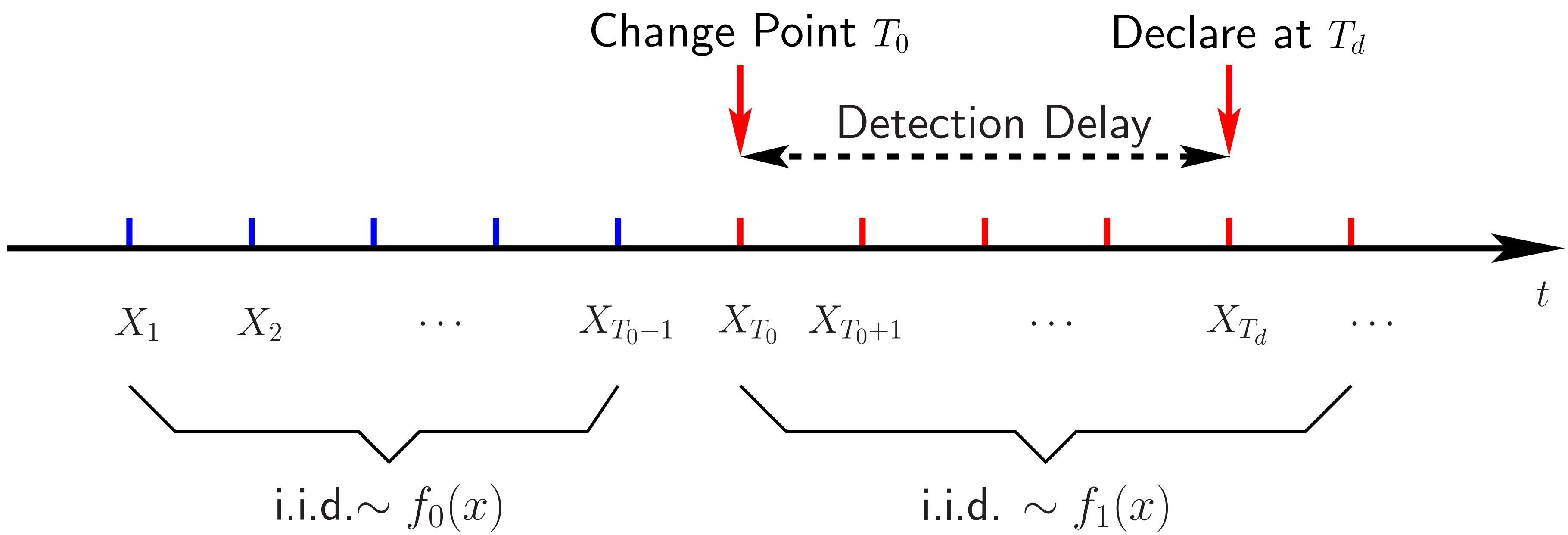
Quickest Search of Spectrum Opportunity

Quickest Search of Spectrum Opportunity



- Sense one channel at a time
- Measurements are taken sequentially.
- Sensing is imperfect.

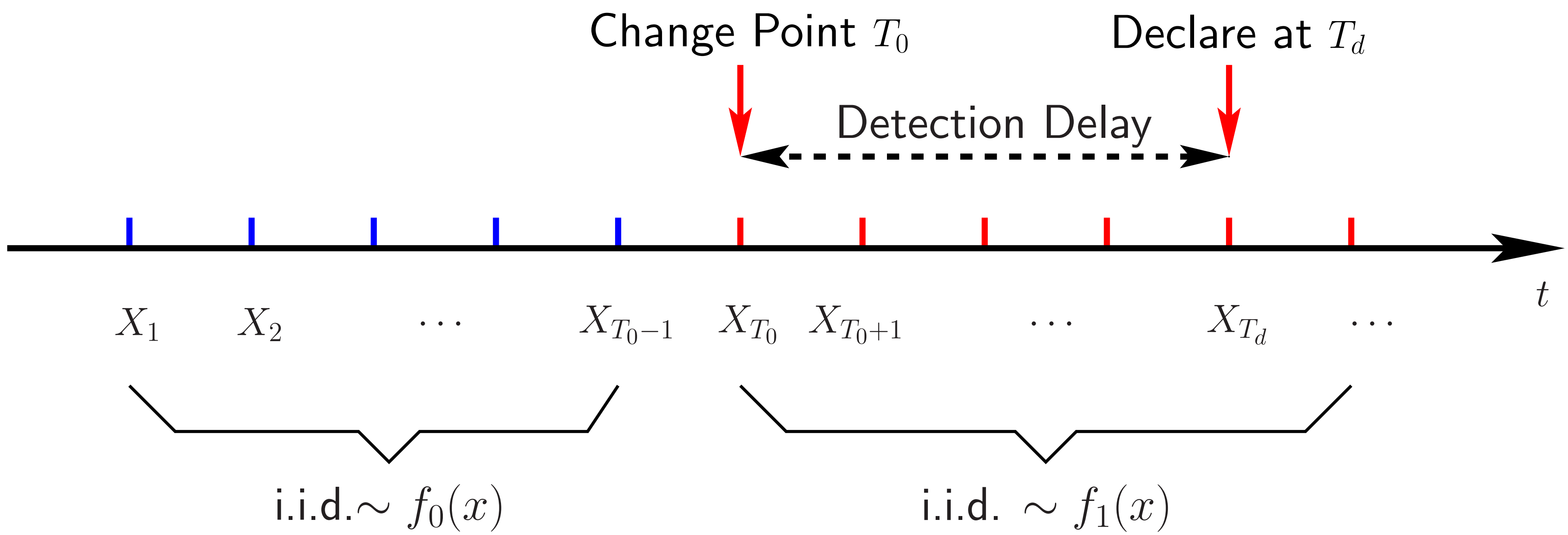
Quickest Change Detection



- Quickest Detection: min $\underbrace{\mathbb{E}[(T_d - T_0)^+]}_{\text{Detection Delay}}$ subject to $\underbrace{\Pr[T_d < T_0] \leq \zeta}_{\text{Reliability Constraint}}$

- Tradeoff: Detection delay vs. detection reliability.

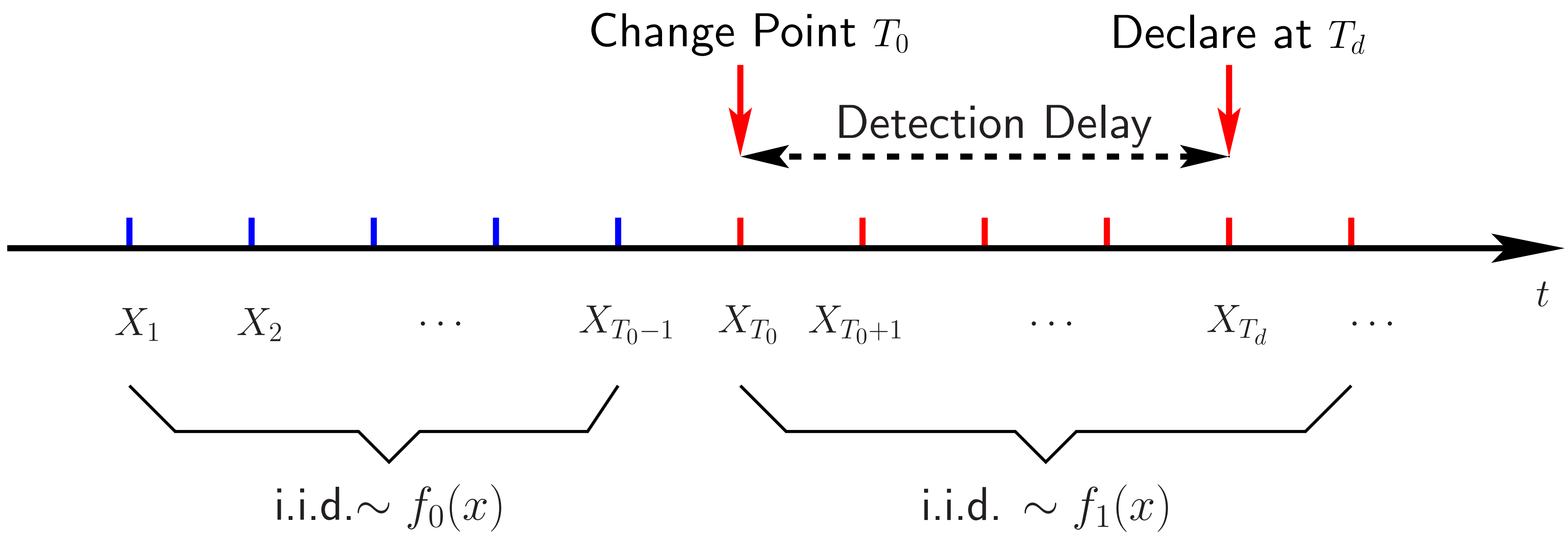
Quickest Change Detection



- **Quickest Detection:** min $\underbrace{\mathbb{E}[(T_d - T_0)^+]}_{\text{Detection Delay}}$ subject to $\underbrace{\Pr[T_d < T_0] \leq \zeta}_{\text{Reliability Constraint}}$

  - Bayesian: Shiryaev'61, Borovkov'98, Tartakovsky&Veeravalli'05.
  - Minimax: CUSUM (Page'54, Lorden'71, Moustakides'86).

Quickest Change Detection: Classic Bayesian Formulation

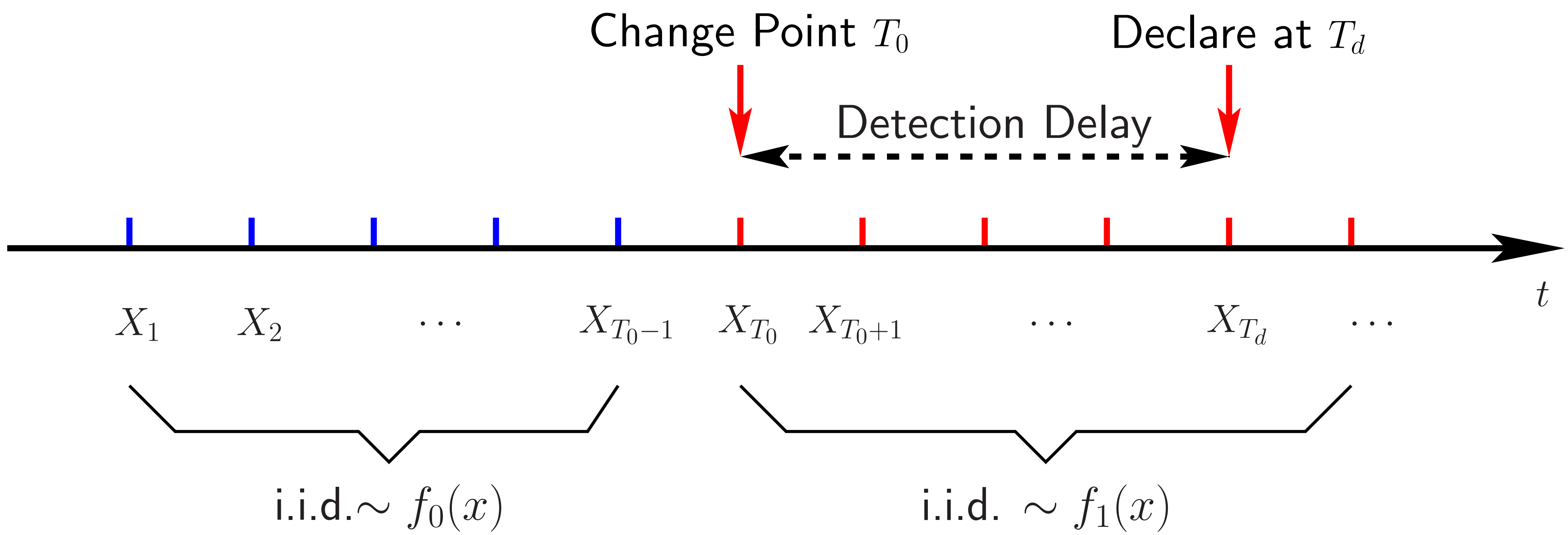


Bayesian Formulation:

- Priori distribution of change point T_0: geometric

$$\begin{aligned}\Pr[T_0 = 0] &= \lambda_0 \\ \Pr[T_0 = k] &= (1-\lambda_0)p(1-p)^{k-1}, \ \forall k > 0,\end{aligned}$$

Shiryaev's Algorithm



- A sufficient statistic: **a posterior probability** that change has occurred

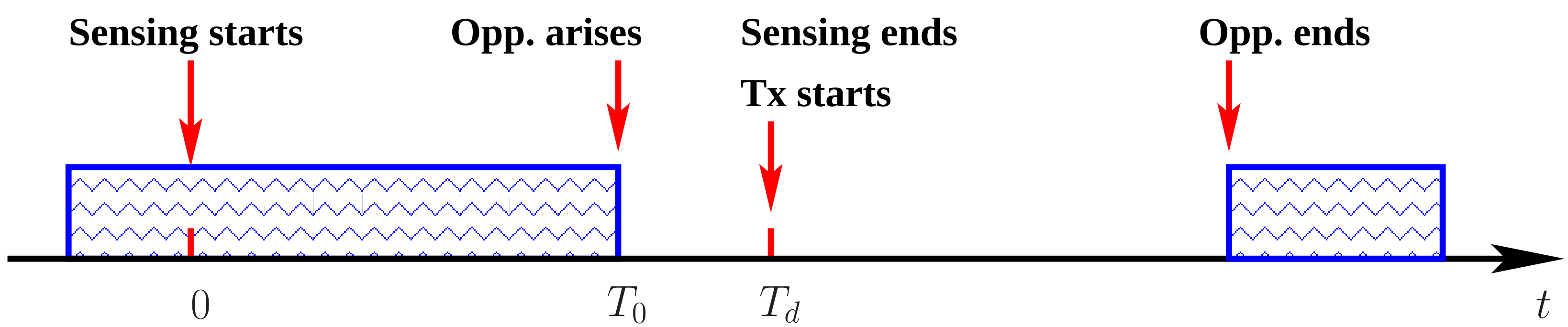
$$\lambda_t \triangleq \Pr[T_0 \leq t | X_1, X_2, \ldots, X_t].$$

- Shiryaev's detection rule:

$$T_d = \inf\{t : \lambda_t \geq \eta_d\}$$

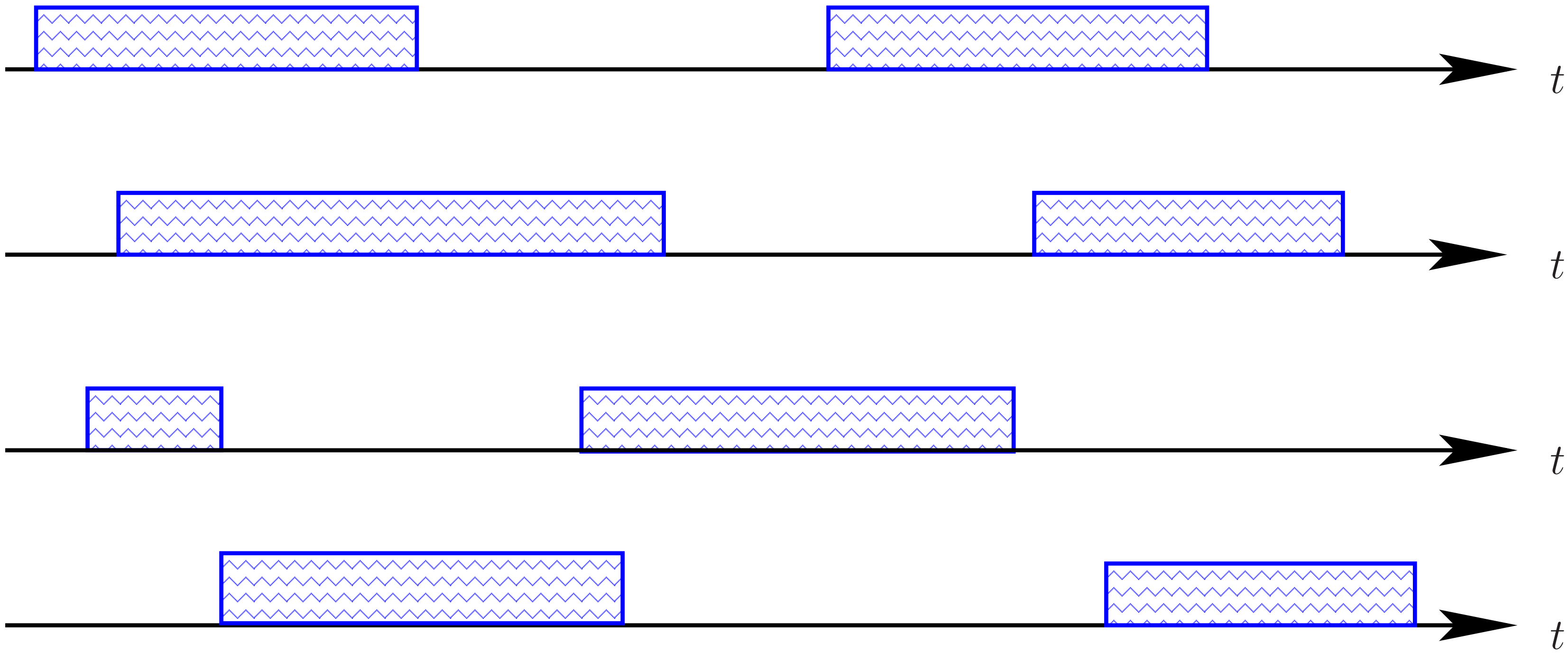
- Detection threshold η_d: determined by the reliability constraint ζ.
- Setting $\eta_d = 1 - \zeta$ is asymptotically optimal as $\zeta \to 0$.

Application in Cognitive Radio



- **Measurements:** $\{X_1, X_2, \ldots, X_{T_0-1}\}$ are i.i.d with distribution $f_0(x)$;
  $\{X_{T_0}, X_{T_0+1}, \ldots\}$ are i.i.d with distribution $f_1(x)$.
- **Stopping Time:** At time $t = T_d$, the user declares an opportunity.
- **Quickest Detection:** min $\underbrace{\mathbb{E}[(T_d - T_0)^+]}_{\text{Detection Delay}}$ subject to $\underbrace{\Pr[T_d < T_0] \leq \zeta}_{\text{Interference Constraint}}$

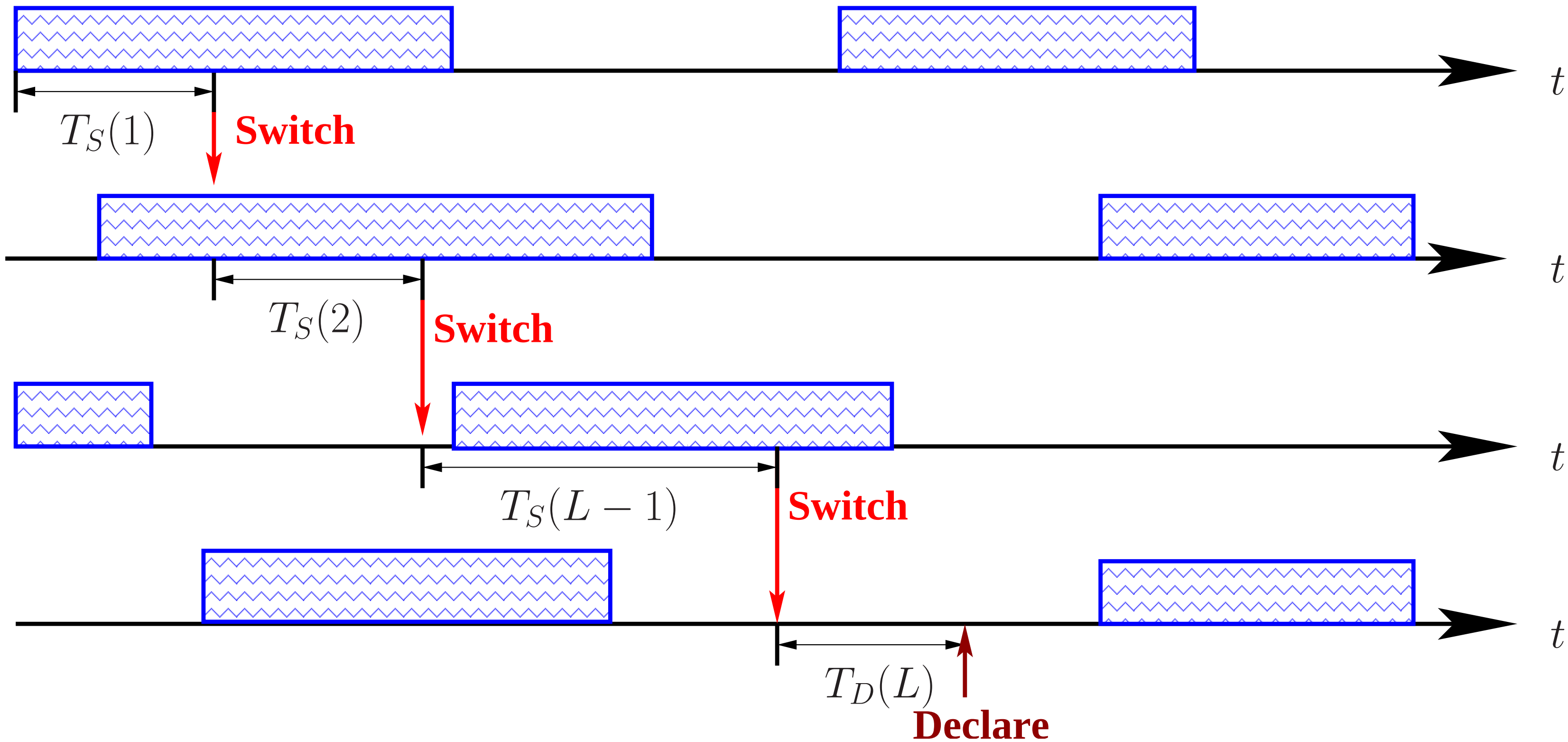
Quickest Search of Opportunity



- Two Fundamental Differences:
  - Channel occupancy is an on-off process with multiple change points.
  - There are multiple channels available.

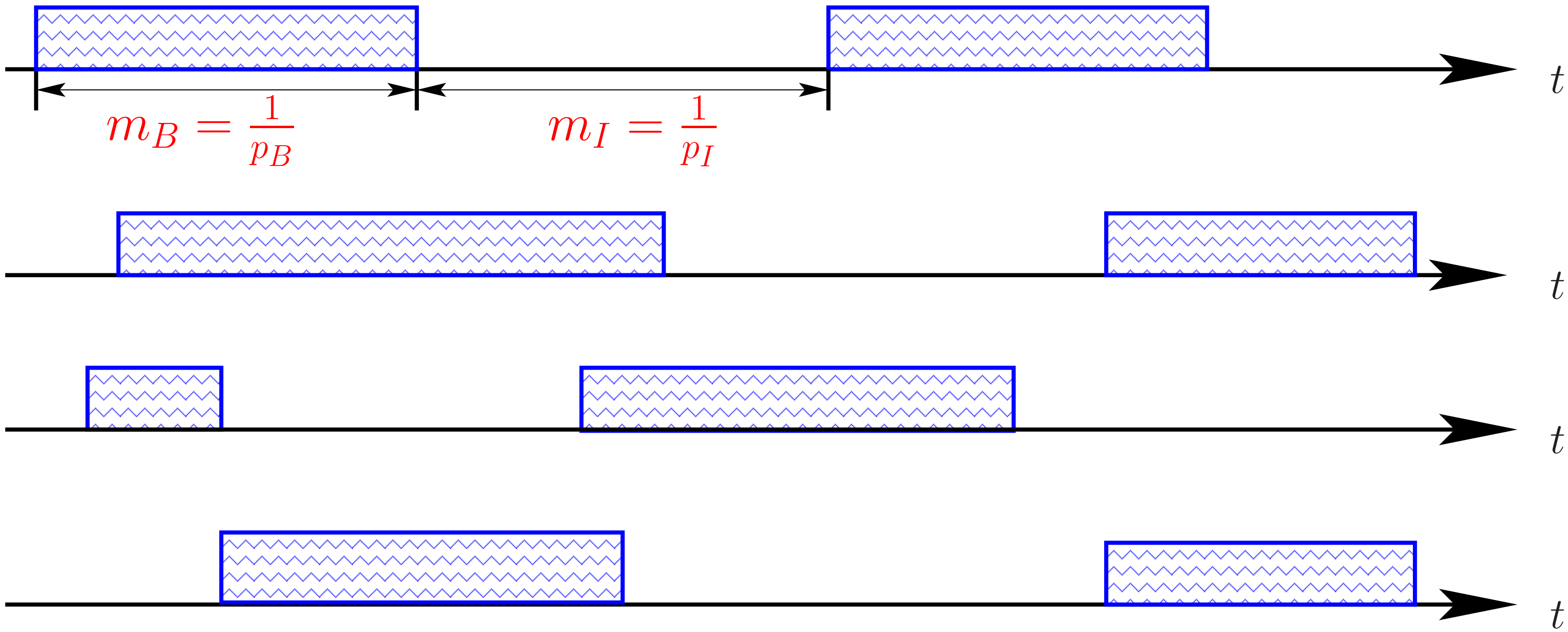
[0] Zhao&Ye'08MILCOM, Zhao&Ye'09ICASSP, Ye&Zhao'09Allerton

Quickest Search of Opportunity



- Quickest Detection of Idle Periods in Multiple On-Off Processes:
  - Continue, switch, or declare?
- Tradeoffs:
  - Whether to declare: delay vs. reliability.
  - Whether to switch: loss of data vs. avoiding bad realizations.

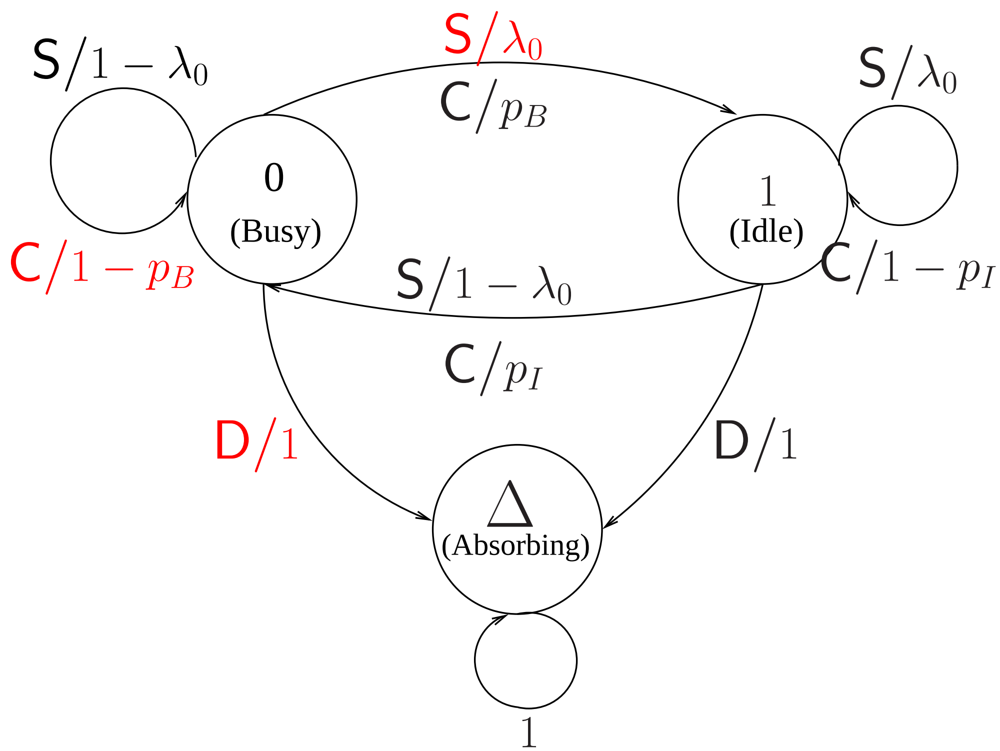
Quickest Search of Opportunity



- A large number of independent homogeneous on-off processes.
- Busy period: geometrically distributed with mean $m_B = \frac{1}{p_B}$.
- Idle period: geometrically distributed with mean $m_I = \frac{1}{p_I}$.
- Fraction of idle time: $\lambda_0 = \frac{m_I}{m_I + m_B}$.

A POMDP Formulation

- State Space: 0 (busy), 1 (idle), $\triangle$ (absorbing state)
- Action Space: S (Switch), C (Continue), D(Declare)
- State Transition:



- Cost:
  - Switch or Continue: 1
  - Declare during a busy period: γ

A POMDP Formulation

- A Sufficient Statistic: the information state (belief)

$$\lambda_t = \Pr[Z_t = \textsf{idle} | X_1, X_2, \ldots, X_t]$$
$$\lambda_0 = \frac{m_I}{m_I + m_B}$$

- Recursive Update of the Information State

$$\lambda_t = \begin{cases} \mathcal{T}(\lambda_0 | x) & a(t-1) = \textsf{S},\ X_t = x \\ \mathcal{T}(\lambda_{t-1} | x) & a(t-1) = \textsf{C},\ X_t = x \end{cases}.$$

- $\mathcal{T}(\lambda|x)$: updated information state based on the new measurement x.

$$\mathcal{T}(\lambda|x) \triangleq \frac{(\lambda \bar{p}_I + \bar{\lambda} p_B) f_1(x)}{(\lambda \bar{p}_I + \bar{\lambda} p_B) f_1(x) + (\lambda p_I + \bar{\lambda} \bar{p}_B) f_0(x)}.$$

A POMDP Formulation

- Search policy π:

$$\lambda_t \in [0,1] \quad \Longrightarrow \quad a(t) \in \{S, C, D\}, \text{ for each time } t.$$

- Quickest Search of Opportunity:

$$\pi^* = \arg\min_{\pi} \mathbb{E}_\pi[\sum_{t=0}^{\infty} \underbrace{R_{\pi(\lambda_t)}}_{\text{Cost}} \mid \lambda_0 = \frac{m_I}{m_B + m_I}],$$

Value Functions

- $V(\lambda_t)$: the minimum expected total cost-to-go when the current belief is λ_t.

$$V(\lambda_t) = \min\{ \underbrace{V_C(\lambda_t)}_{\text{Continue}}, \underbrace{V_S(\lambda_t)}_{\text{Switch}}, \underbrace{V_D(\lambda_t)}_{\text{Declare}} \}.$$

- $V_C(\lambda_t)$: the minimum expected total cost-to-go if continue at t.

$$V_C(\lambda_t) = 1 + \int_x \underbrace{P(x;\lambda_t)}_{\Pr[\text{ observe } x \text{ under } \lambda_t\ ]} V(\mathcal{T}(\lambda_t|x))\mathrm{d}x$$

- $V_S(\lambda_t)$: the minimum expected total cost-to-go if switch at t.

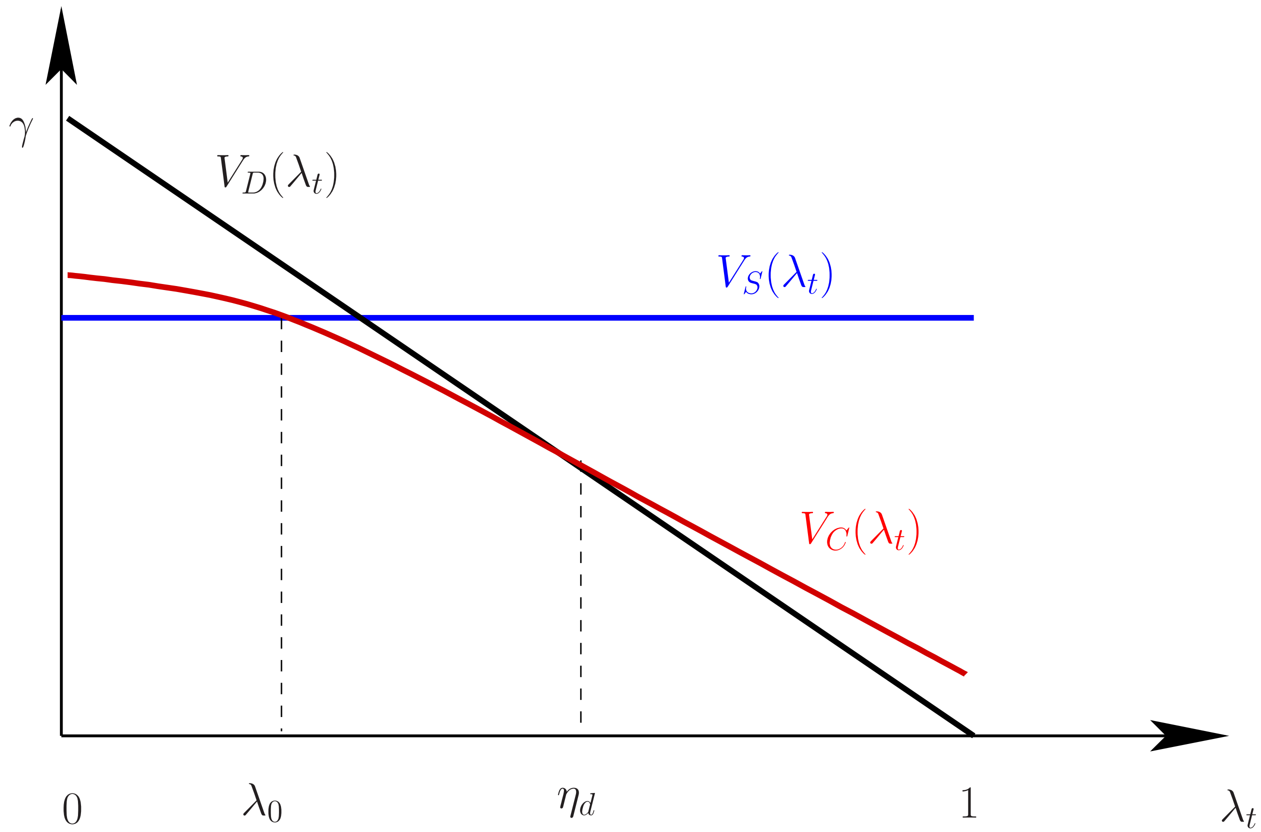
$$V_S(\lambda_t) = 1 + \int_x \underbrace{P(x;\lambda_0)}_{\Pr[\text{ observe } x \text{ under } \lambda_0\ ]} V(\mathcal{T}(\lambda_0|x))\mathrm{d}x = V_C(\lambda_0)$$

- $V_D(\lambda_t)$: the minimum expected total cost-to-go if declare at t.

$$V_D(\lambda_t) = (1 - \lambda_t)\gamma.$$

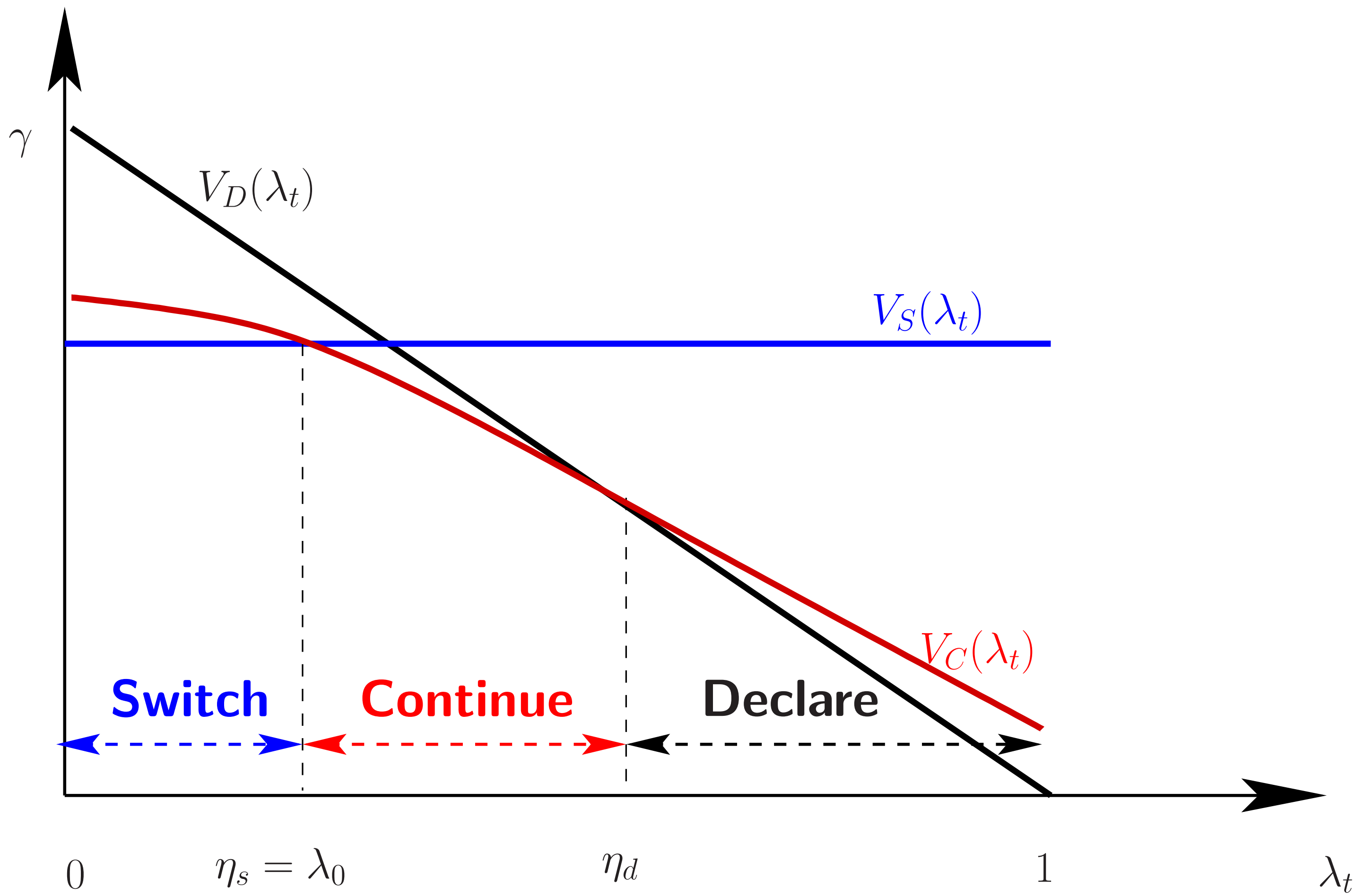
The Optimality of A Threshold Policiy

- $V_D(\lambda_t)$ is linear.
- $V_C(\lambda_t)$ is monotonically decreasing and concave.
- $V_S(\lambda_t) = V_C(\lambda_0)$, where $\lambda_0 = \frac{m_I}{m_I + m_B}$.



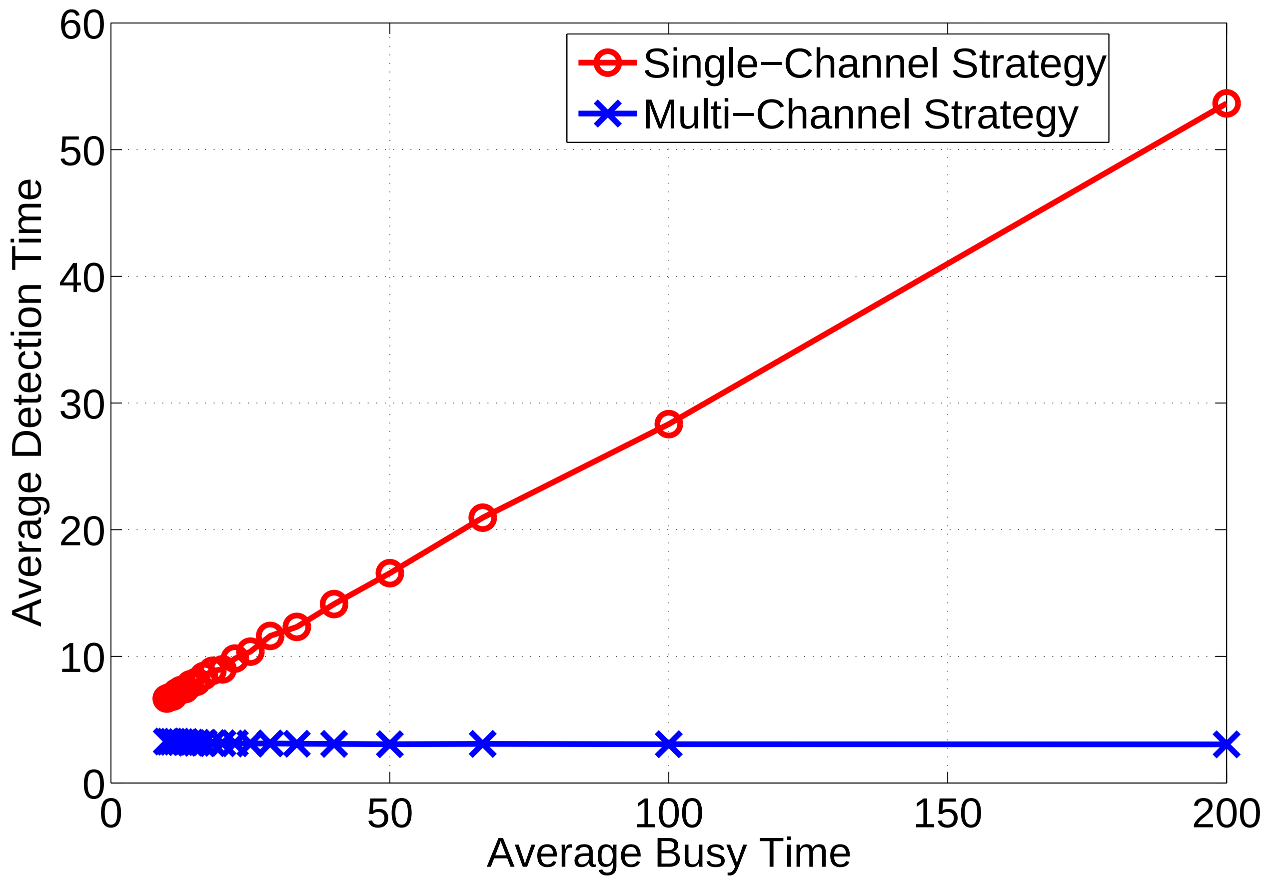
The Optimality of A Threshold Policiy

$$V(\lambda_t) = \min\{ \underbrace{V_S(\lambda_t)}_{\text{Switch}}, \underbrace{V_C(\lambda_t)}_{\text{Continue}}, \underbrace{V_D(\lambda_t)}_{\text{Declare}} \}.$$



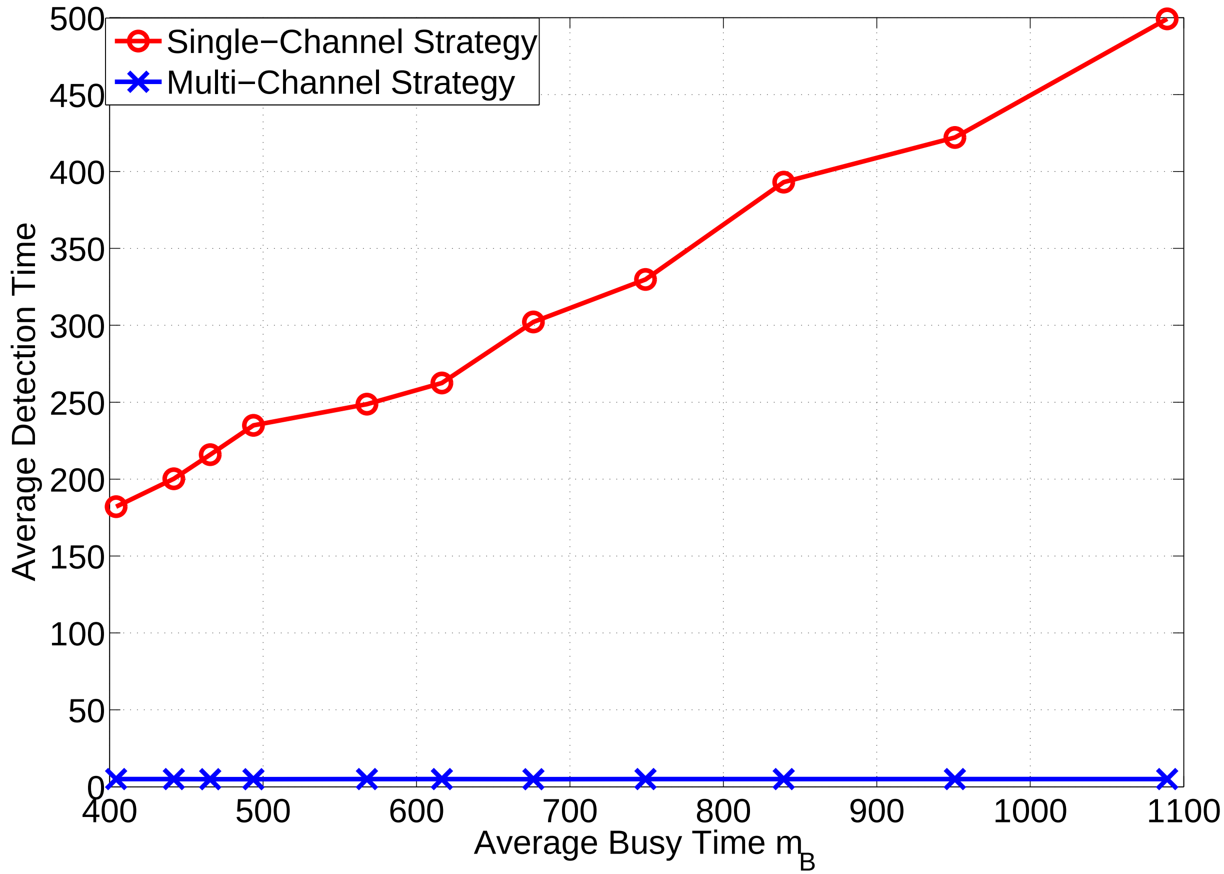
Simulation Example: Geometric Distribution

- Increase both m_B and m_I while keeping λ_0 fixed



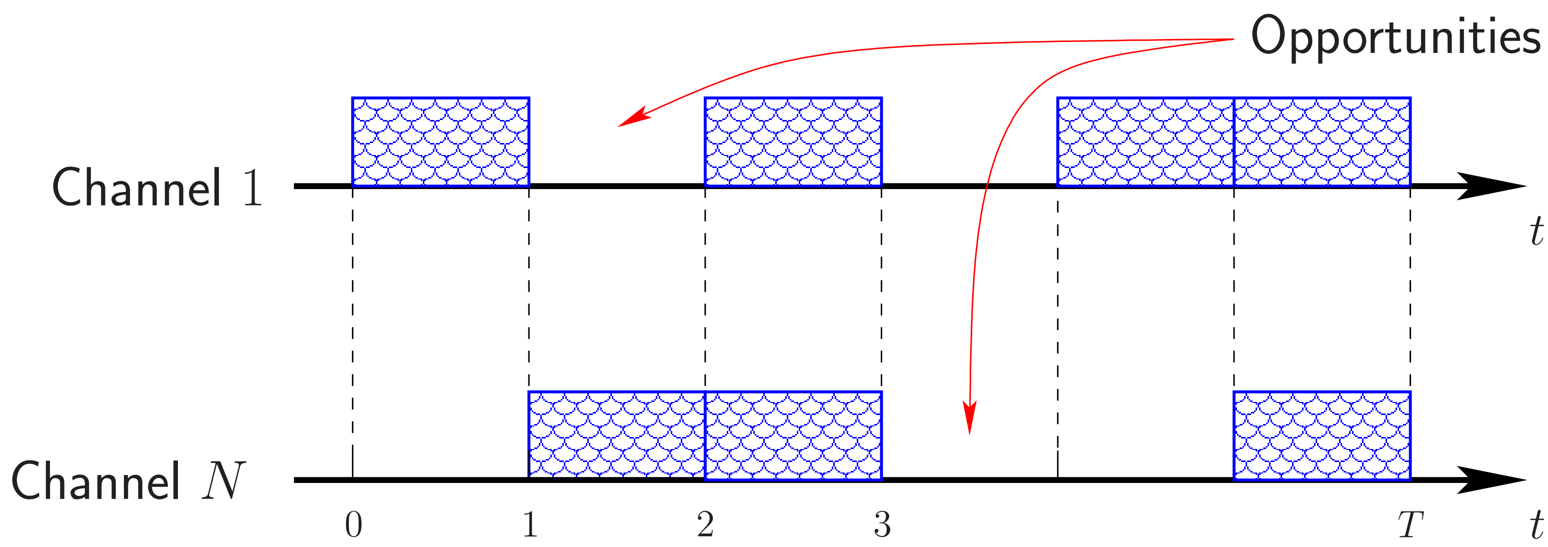
Simulation Example: Arbitrary Distributions

- Busy period: Pareto distribution with increasing tail index



Distributed Learning for Spectrum Sharing under Unknown Model

Cognitive Radio Networks



- N channels, M ($M < N$) **distributed** secondary users (no info exchange).
- Primary occupancy of channel i: i.i.d. Bernoulli with unknown mean θ_i:

$$\theta_i = \Pr[\text{idle}] \underbrace{\Pr[\text{correct detection} \mid \text{idle}]}_{\textit{sensing error incorporated}}$$

- Accessing an idle channel results in a unit reward.
- Users accessing the same channel collide; no one or only one receives reward.
- Objective: decentralized policy for optimal **network-level** performance.

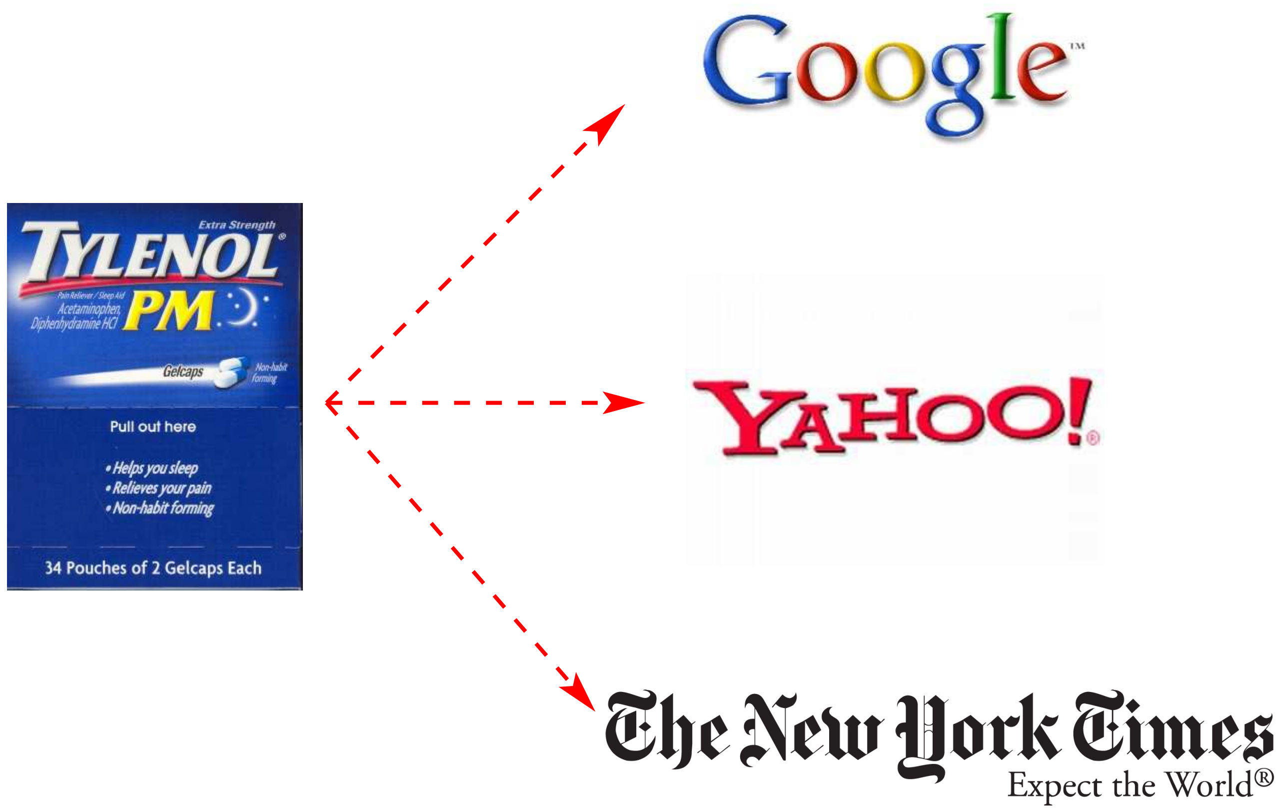
Classic Multi-Armed Bandit

Clinical Trial *(Thompson'33)*

Two treatments with unknown effectiveness:

Web Search and Internet Advertising

Where to place ads?



An Example: Bernoulli Reward

A Two-Armed Bandit:

- Two coins with unknown bias θ_1, θ_2.
- Head: reward $= 1$; Tail: reward $= 0$.
- Objective: maximize long-term total reward.

Non-Bayesian Formulation

- (θ_1, θ_2) are treated as unknown deterministic parameters.
- $V_T^\pi(\theta_1, \theta_2)$: total reward of policy π over a horizon of length T.
- $T \underbrace{\max\{\theta_1, \theta_2\}}_{\theta_{max}}$: total reward if (θ_1, θ_2) were known.
- The cost of learning (regret):

$$R_T^\pi(\theta_1, \theta_2) \triangleq T\theta_{max} - V_T^\pi(\theta_1, \theta_2) = (\theta_{max} - \theta_{min})\mathbb{E}[\text{time spent on } \theta_{min}]$$

- Objective: minimize the rate that $R_T^\pi(\theta_1, \theta_2)$ grows with T.

Classic Results

- Lai&Robbins'85:

$$R_T^*(\theta_1, \theta_2) \sim \underbrace{\frac{\theta_{max} - \theta_{min}}{I(\theta_{min}, \theta_{max})}}_{\textit{KL distance}} \log T \quad \text{as } T \to \infty$$

- Anantharam&Varaiya&Walrand'87:
  - extension from single play to multiple plays.

Classic and Recent Results

- Lai&Robbins'85:

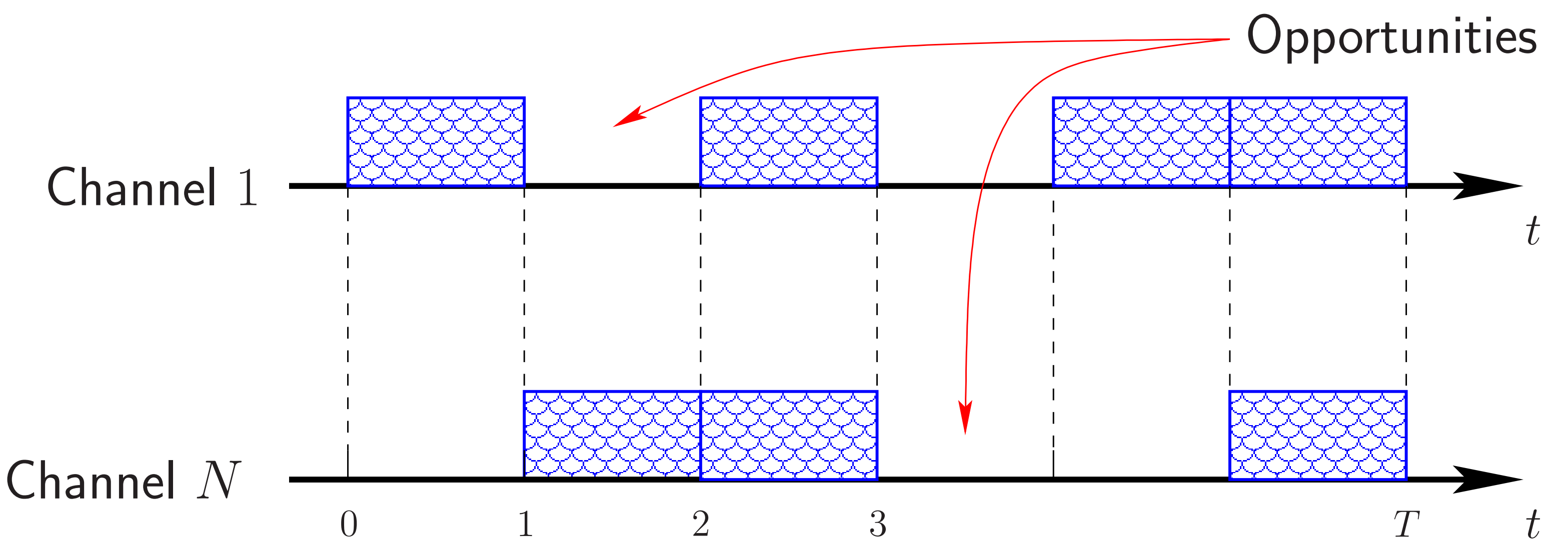
$$R_T^*(\theta_1,\theta_2) \sim \underbrace{\frac{\theta_{max}-\theta_{min}}{I(\theta_{min},\theta_{max})}}_{\textit{KL distance}} \log T \quad \text{as } T \to \infty$$

- Anantharam&Varaiya&Walrand'87:
  - extension from single play to multiple plays.

- Liu&Zhao'10:
  - extension to **distributed** multiple players *(distributed decision-making using only local observations)*.
  - decentralized policy achieving **the same** $\log T$ order of the regret.
  - fairness among players.

Decentralized Multi-Armed Bandit

Decentralized Multi-Armed Bandit



- N arms with **unknown** reward statistics $(\theta_1, \cdots, \theta_N)$.
- M $(M < N)$ distributed players.
- Each player selects one arm to play and observes the reward.
- Distributed decision making using only **local** observations.
- Colliding players either share the reward or receive no reward.

---

[0]Liu&Zhao'10ITA, Liu&Zhao'10ICASSP, Liu&Zhao'10TSP

Decentralized Multi-Armed Bandit

System Regret:

- $V_T^\pi(\Theta)$: total system reward under a decentralized policy π.
- Total system reward with **known** $(\theta_1, \cdots, \theta_N)$ and **centralized scheduling**:

$$T \; \Sigma_{i=1}^{M} \underbrace{\theta^{(i)}}_{i\text{th best}}$$

- System regret:

$$R_T^\pi(\Theta) = T\Sigma_{i=1}^{M}\theta^{(i)} - V_T^\pi(\Theta)$$

The Minimum Regret Growth Rate

The Minimum regret rate in Decentralized MAB is logarithmic.

$$R_T^*(\Theta) \sim C(\Theta) \log T$$

The Key to Achieving $\log T$ Order:

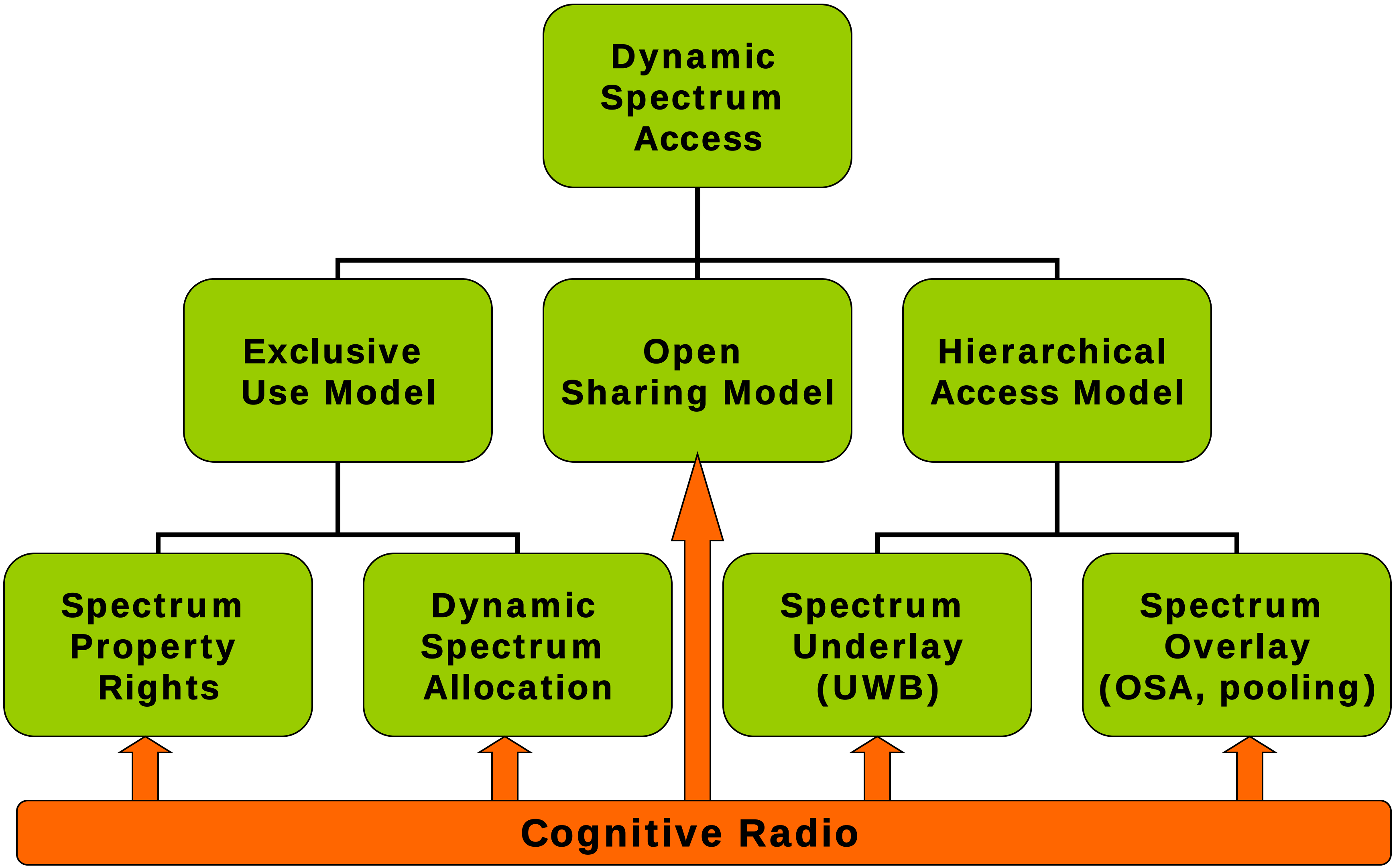
- Learn from local observations which channels are the most rewarding.
- Learn from **collisions** to achieve efficient sharing with other users.

A Framework for Constructing Decentralized Policies

The TDFS (Time Division Fair Sharing) Framework:

- Players use time sharing of the M best arms with different offset.
- Each player learns the M best arms based on local observations.
- Players **learn from collisions** to settle at different offset.
- No pre-agreement or a global time required.

Conclusion



Dynamic Spectrum Access
Exclusive Use Model
Open Sharing Model
Hierarchical Access Model
Spectrum Property Rights
Dynamic Spectrum Allocation
Spectrum Underlay (UWB)
Spectrum Overlay (OSA, pooling)
Cognitive Radio