Centrum Wiskunde & Informatica

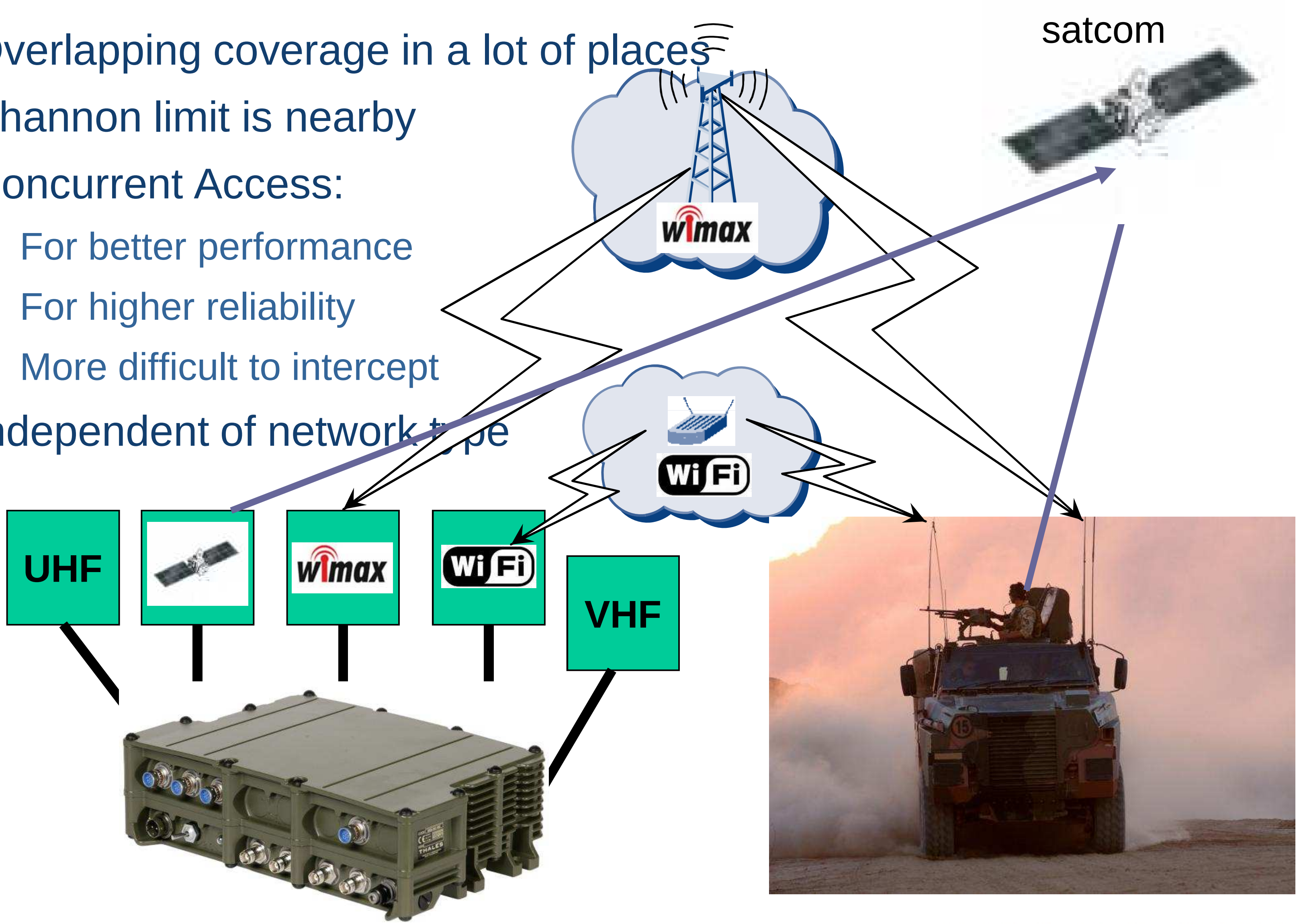
Dynamic Traffic Splitting to Parallel Wireless Networks with Partial Information: A Bayesian Approach

Gerard Hoekstra (Thales/CWI), **gerard.hoekstra@nl.thalesgroup.com**
with Sandjai Bhulai (VU University Amsterdam / CWI) and
Rob van der Mei (CWI / VU University Amsterdam)

YEQT-IV (Young European Queueing Theorists)
"Optimal Control in Stochastic Systems", November 25-27, 2010

Motivation: Using all available networks

- Overlapping coverage in a lot of places
- Shannon limit is nearby
- Concurrent Access:
  - For better performance
  - For higher reliability
  - More difficult to intercept
- Independent of network type



Limits on Increasing Performance of Wireless

- Currently: in practical communication systems the theoretical channel capacity (Shannon Limit) is approached.

IEEE Communications Magazine • December 2008

FUNDAMENTAL LIMITATIONS ON INCREASING DATA RATE IN WIRELESS SYSTEMS

DONALD C. COX

HAROLD TRAP FRISS PROFESSOR OF ELECTRICAL ENGINEERING, STANFORD UNIVERSITY

ISSUE 11 Newsletter of Institute for Mathematical Sciences, NUS 2007

Mathematical Conversations

Sergio Verdú: Wireless Communications, at the Shannon Limit >>>

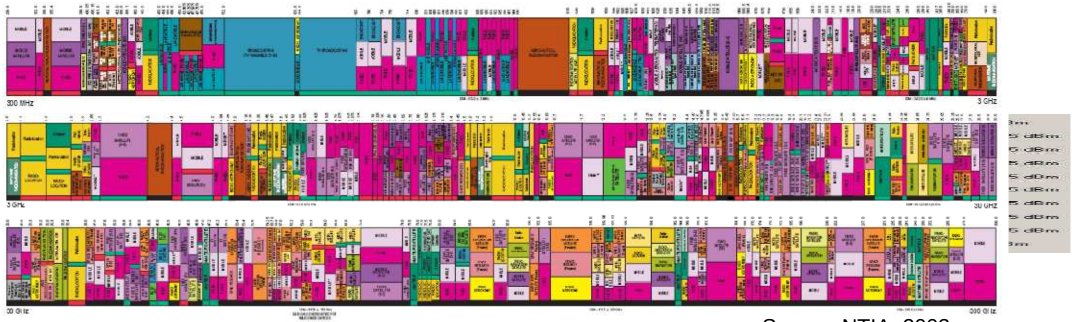
Satellites Approach Theoretical Shannon Limit

ScienceDaily (Nov. 3, 2008) — Satellites are achieving unparalleled efficiency with a new protocol, DVB-S2. The performance of DVB-S2 satellite systems is very close to the theoretical maximum, defined by the Shannon Limit. That efficiency could be pushed even further by network optimisation tools and equipment recently developed by European researchers.

- Future: modest increases expected in data transmission rates from sophisticated signal processing (e.g. Multiple Input Multiple Output).

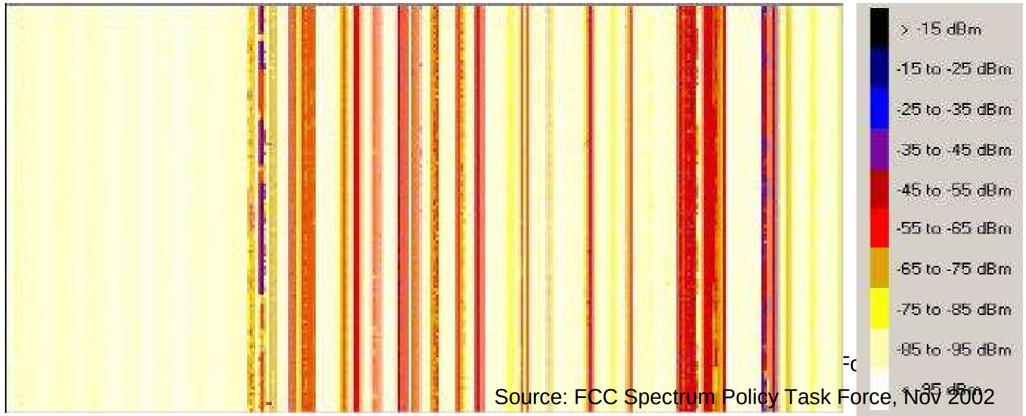
Efficiency over a larger spectrum

- So, high spectrum efficiency of wireless networking technologies operating *within* the different allocated frequency bands (e.g. SatCom, WiFi).



Source: NTIA, 2003

- Measurements of actual spectrum usage reveals *idle* bands in the seemingly crowded RF spectrum

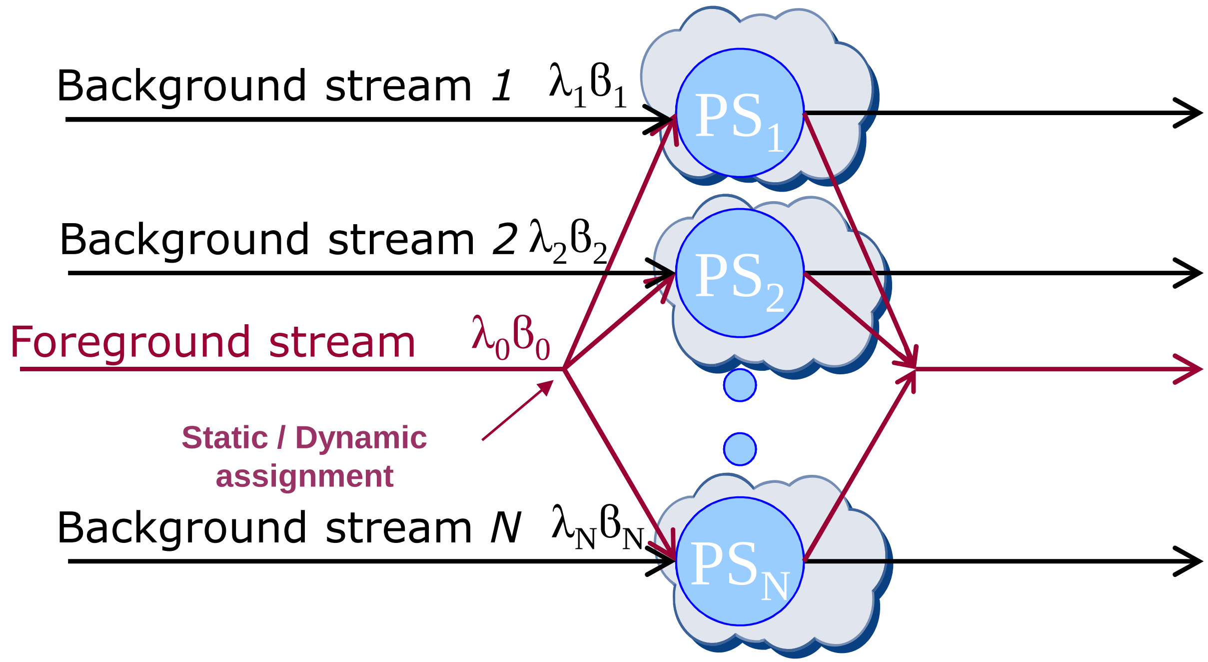


Source: FCC Spectrum Policy Task Force, Nov 2002

Spectrum isn't scarce

How to minimize the mean number of foreground jobs in the system of N parallel PS-nodes

(in the presence of background traffic) ?

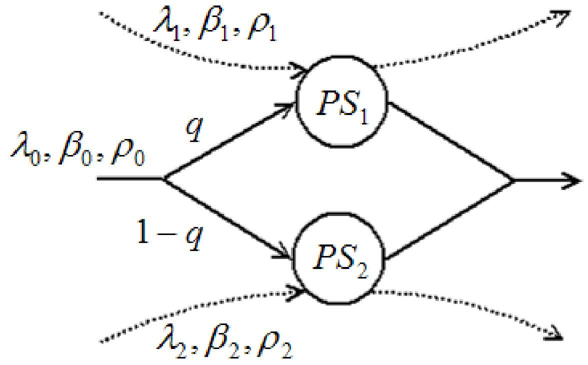


Assumptions:

- Poisson arrivals
- General service times
- Foreground and background traffic

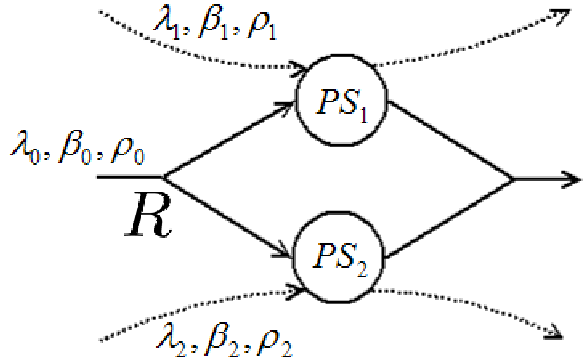
1. **Static assignment**
   Minimizing mean number of foreground jobs in the system:

$$E[N] = \frac{q}{1-q\rho_0-\rho_1} + \frac{1-q}{1-(1-q)\rho_0-\rho_2}$$



2. **Dynamic assignment**
   Calculate optimal decision policy:

$$\begin{aligned} TV(s) = & \sum_{i=1}^{k} x_i + \lambda_0 \min_{a \in \{1,\dots,k\}} \{V(s+e_a)\} + \\ & \sum_{i=1}^{k} \lambda_i V(s+e_{i+k}) + \sum_{i=1}^{k} \frac{x_i}{x_i+y_i} \mu_0 V(s-e_i) + \\ & \sum_{i=1}^{k} \frac{y_i}{x_i+y_i} \mu_i V(s-e_{i+k}) + \\ & \Big(1 - \lambda_0 - \sum_{i=1}^{k} \big[\lambda_i + \frac{x_i}{x_i+y_i}\mu_0 + \frac{y_i}{x_i+y_i}\mu_i\big]\Big) V(s) \end{aligned}$$



Full Observability model

(Exp. Service times)

Partial Observability for *N* Networks

- In practice only total number of jobs in a network may be observed

- Model statespace $(x_1, \ldots, x_N, y_1, \ldots, y_N)$

- Decisions based on $(z_1, \ldots, z_N)$

Bayesian Partial Information Model

- Bayesian policy maps $(z_1, \ldots, z_N) \rightarrow (x_1, \ldots, x_N, y_1, \ldots, y_N)$
- Mapping must account for the complete history of states
- State space of the Bayesian program consists of the observation $z = (z_1, \ldots, z_N) \in \mathbb{N}_0^N$ and information state $\prod_{i=1}^N \{u_i \in [0,1]^{\mathbb{N}_0} \mid \sum_{x \in \mathbb{N}_0} u_i(x) = 1\}$.

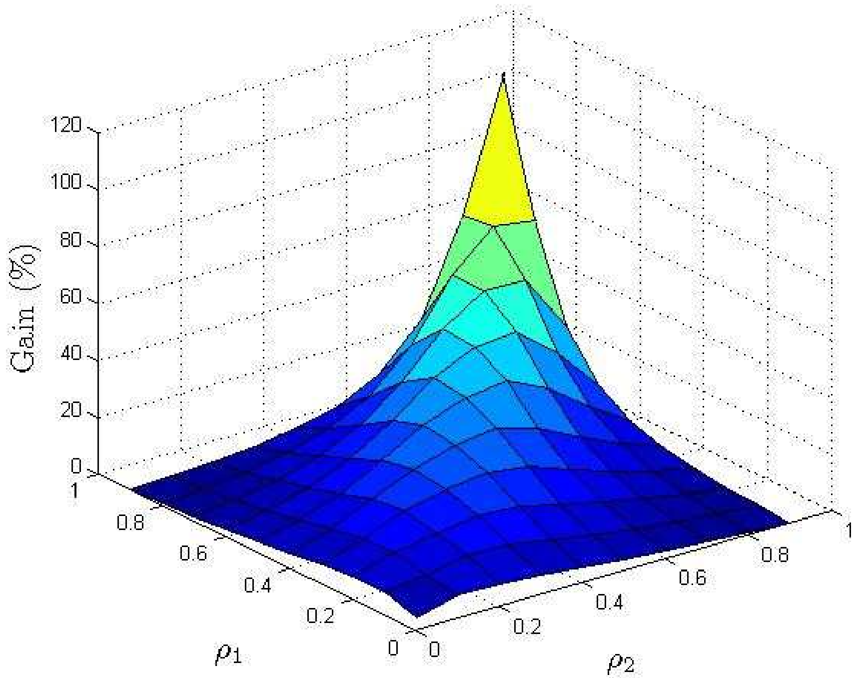
- Every job departure and arrival provides information
- Departure types cannot be observed, → apply information state.

$$
\begin{aligned}
TV(s) = \sum_{x_1 \in \mathbb{N}_0} \cdots \sum_{x_N \in \mathbb{N}_0} u_1(x_1) \cdots u_N(x_N) \Big[& \sum_{i=1}^N x_i + \sum_{i=1}^N \lambda_i V(s_{ab_i}) + \\
& \lambda_0 \min\{V(s_{af_1}), \ldots, V(s_{af_N})\} + \\
& \sum_{i=1}^N \tfrac{x_i}{z_i} \mu_0 V(s_{df_i}) + \sum_{i=1}^N \tfrac{z_i - x_i}{z_i} \mu_i V(s_{db_i}) + \\
& \Big(1 - \lambda_0 - \sum_{i=1}^N \big[\lambda_i + \tfrac{x_i}{z_i}\mu_0 + \tfrac{z_i - x_i}{z_i}\mu_i\big]\Big) V(s) \Big].
\end{aligned}
$$

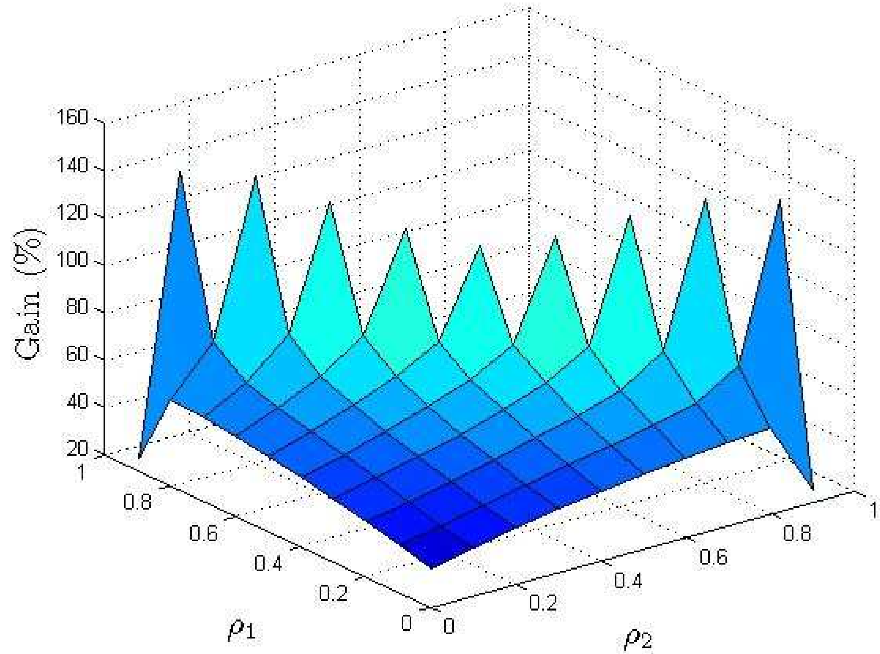
THALES

Comparison of Static vs Dynamic Assignment

- Foreground performance gain $\frac{\mathbb{E}N_{\text{static}} - \mathbb{E}N_{\text{dynamic}}}{\mathbb{E}N_{\text{dynamic}}} \cdot 100\%$ of dynamic (full observability) over static policy:
  - Large when background traffic loads are equal (Light)
  - Decreasing gains as the background traffic load differences increase (Light)
  - Highly sensitive to decision subtilty (non-Light)



Light foreground traffic load (ρ_0=0.1)



Medium foreground traffic load (ρ_0=0.9)

THALES

- Simulating dynamic policies in equal capacity PS-nodes

$$\left(\mathbb{E}\left[S|\text{Bayes}\right], \mathbb{E}\left[S|\text{full MDP}\right], \Delta\%\right)$$

$\beta_0 = \beta_1 = \beta_2 = 1$

Background traffic load on node 1

$\rho_1 \backslash \rho_2$	0.1	0.3	0.5	0.7	0.9
0.1	(1.073, 1.072, 0.2%)	(1.116, 1.115, 0.1%)	(1.165, 1.164, 0.1%)	(1.210, 1.210, 0.0%)	(1.245, 1.241, 0.4%)
0.2		(1.196, 1.195, 0.1%)	(1.273, 1.271, 0.2%)	(1.352, 1.350, 0.1%)	(1.416, 1.409, 0.5%)
0.3		(1.282, 1.277, 0.4%)	(1.393, 1.390, 0.2%)	(1.516, 1.514, 0.1%)	(1.628, 1.625, 0.2%)
0.4			(1.522, 1.519, 0.2%)	(1.715, 1.711, 0.2%)	(1.918, 1.911, 0.3%)
0.5			(1.674, 1.665, 0.5%)	(1.958, 1.952, 0.3%)	(2.318, 2.308, 0.4%)
0.6				(2.260, 2.255, 0.3%)	(2.910, 2.897, 0.5%)
0.7				(2.654, 2.641, 0.5%)	(3.878, 3.858, 0.5%)
0.8					(5.735, 5.705, 0.5%)
0.9					(11.064, 11.020, 0.4%)

$\rho_0 = 0.1$

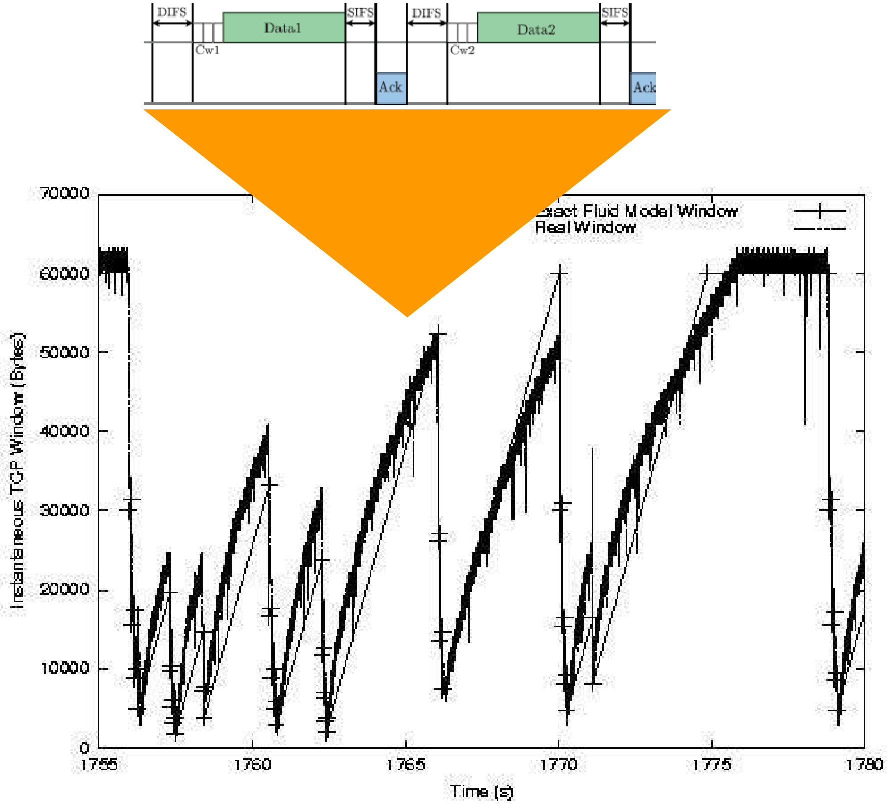
Background traffic load on node 2

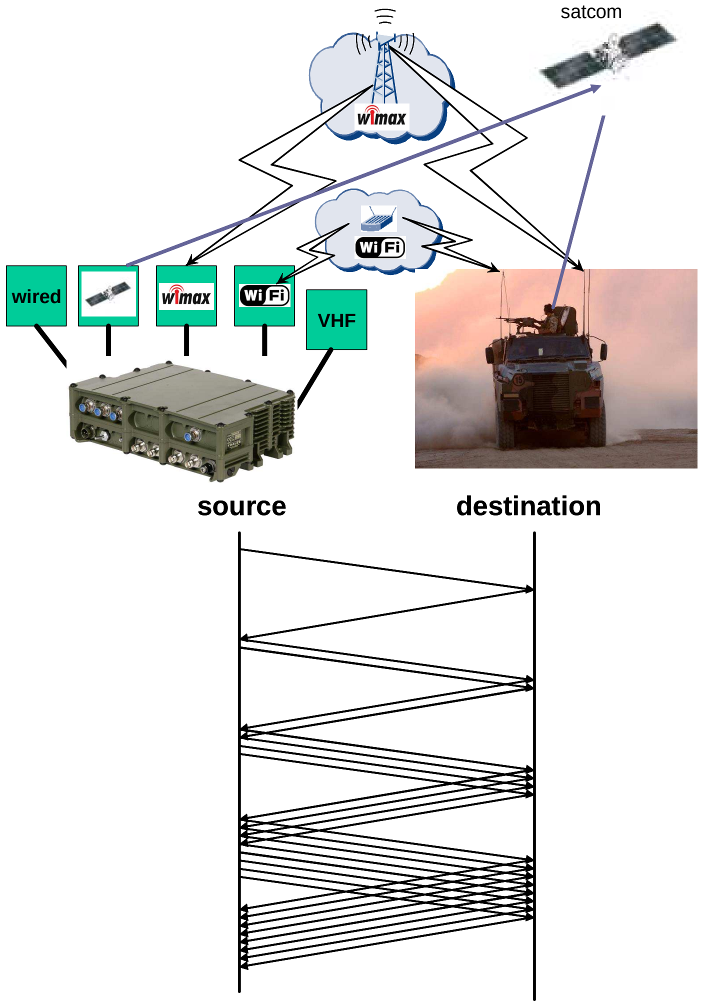
$\rho_1 \backslash \rho_2$	0.1	0.3	0.5	0.7	0.9
0.1	(1.59, 1.55, 2.3%)	(1.93, 1.88, 2.6%)	(2.53, 2.46, 2.9%)	(3.76, 3.70, 1.7%)	(8.90, 8.89, 0.2%)
0.2		(2.21, 2.15, 2.9%)	(3.08, 2.99, 2.8%)	(5.31, 5.23, 1.5%)	unstable
0.3		(2.60, 2.51, 3.5%)	(3.99, 3.86, 3.4%)	(9.74, 9.63, 1.1%)	unstable
0.4			(5.67, 5.60, 1.3%)	unstable	unstable
0.5			(10.87, 10.62, 2.4%)	unstable	unstable

$\rho_0 = 0.9$

- Bayesian algorithm performs close to full-obs.

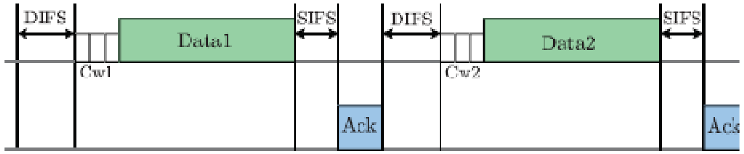
Applicaton to communication networks



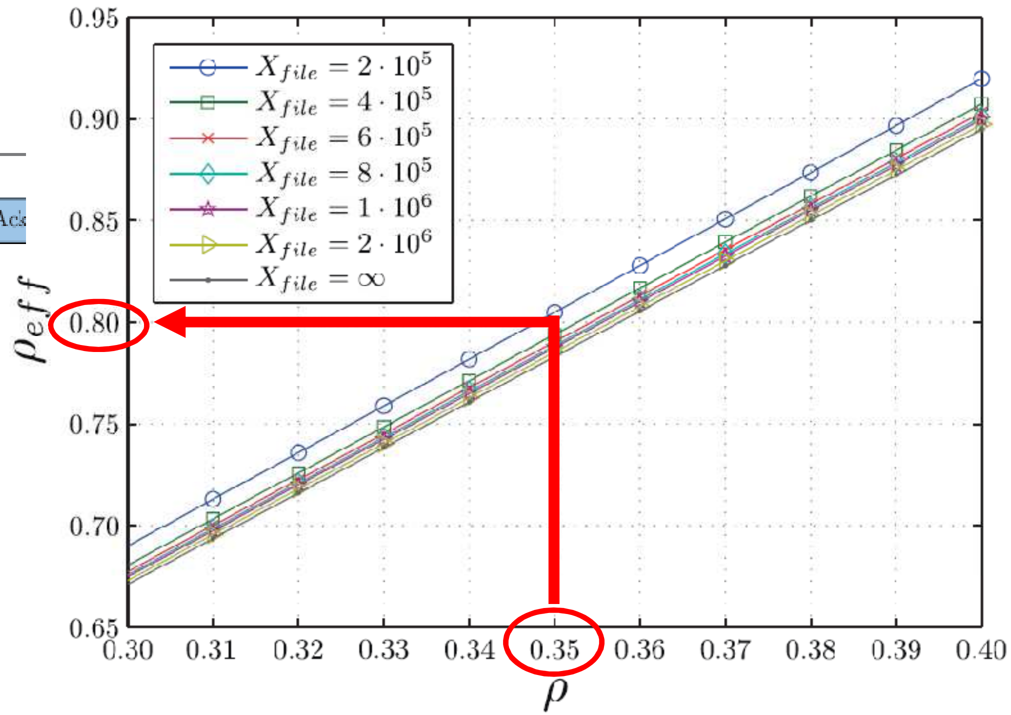


- Practical complexity:
  - Packet level dynamics
  - MAC transmission cycles
  - Protocol interplay

THALES



Load of 35% may generate easily 80% effective load



- Effective service rate of a transfer depends on many lower-level details (TCP parameters, MAC parameters PHY...)

“Effective load”: $\rho_{eff} := \lambda \dfrac{X_{file}}{TP_{effective}}$

average file size

“effective throughput”

- Dynamic policies in equal capacity networks

$$\left(\mathbb{E}\left[S|\text{Bayes}\right], \mathbb{E}\left[S|\text{full MDP}\right], \Delta\%\right)$$

Background traffic load on network 1

$\rho_1 \backslash \rho_2$	0.1	0.3	0.5	0.7	0.8
0.1	(0.365, 0.354, 0.31%)	(0.370, 0.369, 0.30%)	(0.385, 0.385, 0.05%)	(0.402, 0.400, 0.49%)	(0.408, 0.406, 0.46%)
0.2		(0.396, 0.395, 0.21%)	(0.420, 0.420, 0.10%)	(0.446, 0.446, 0.05%)	(0.456, 0.455, 0.11%)
0.3		(0.421, 0.421, 0.18%)	(0.460, 0.458, 0.33%)	(0.502, 0.498, 0.65%)	(0.521, 0.519, 0.48%)
0.4			(0.503, 0.501, 0.35%)	(0.565, 0.564, 0.18%)	(0.601, 0.599, 0.37%)
0.5			(0.551, 0.547, 0.61%)	(0.643, 0.639, 0.68%)	(0.700, 0.696, 0.53%)
0.6				(0.742, 0.737, 0.68%)	(0.831, 0.828, 0.29%)
0.7				(0.867, 0.865, 0.20%)	(1.028, 1.018, 1.01%)
0.8					(1.322, 1.319, 0.23%)

Background traffic load on network 2

$\rho_0 = 0.1$

- Bayesian algorithm again closely matches full observation in real networking environment.

(Traffic loads are “Effective Loads”)

Remark:

- Bayesian updates in this case (due to the arrival type observability→ deterministic state transition) keeps dimensionality at reasonably low levels.

Challenges:

- Performance of the Bayesian policy in a network with unequal service capacity.
- Impact of policies on background traffic performance.
- Phase-type distributions for jobs