QoS-Driven Optimal Resource Management

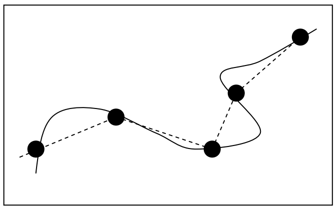
Chang-Gun Lee
Seoul National University

Resource Assignment Problem

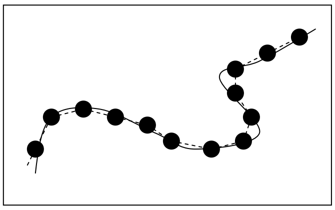
- Resources: CPU, Bandwidth, etc.
- Resource Assignment Problem: *How much resource should be assigned to real-time tasks?*
  - Always satisfy minimal timing constraints (or QoS)
    - Minimal resource assignment
  - Optimize QoS with left over resource
    - More resource → Better Quality

Quality vs. Resource

- Target Tracking Quality
  - Minimal sampling rate for the minimal quality
  - Better tracking quality by using higher sampling rate and more sophisticated algorithms (allocate more CPU cycles).



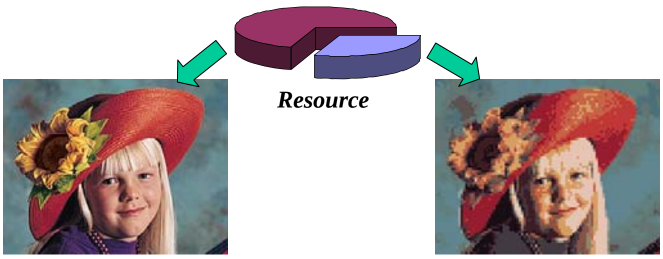
Tracking with a low sampling rate



Tracking with a high sampling rate

Quality vs. Resource

- Quality vs. Invested Amount of Resource
- Optimal Use of Limited Resource for Best Quality

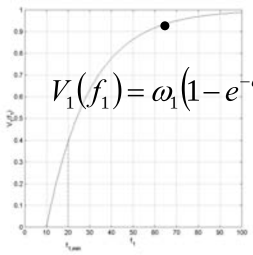


Resource Assignment Problem

- How *multiple tasks* share the *limited resource* such that the **sum of quality can be maximized**?

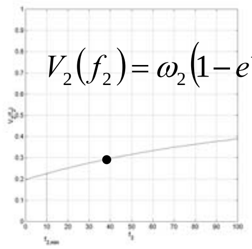
Optimal Frequency Assignment for Quality Optimization

Missile



$$V_1(f_1) = \omega_1\left(1 - e^{-\alpha_1 f_1 + \beta_1}\right)$$

Air plane



$$V_2(f_2) = \omega_2\left(1 - e^{-\alpha_2 f_2 + \beta_2}\right)$$

Maximize $V_1(f_1) + V_2(f_2)$

Subject to $c_1 f_1 + c_2 f_2 \le UB$

$f_1 \ge f_{1,\min}, f_2 \ge f_{2,\min}$

Problem Formulation

Maximize $\sum_{i=1}^{n} V_i(f_i) = \sum_{i=1}^{n} \omega_i \left(1 - e^{-\alpha_i f_i + \beta_i}\right)$

Subject to $\sum_{i=1}^{n} c_i f_i \le UB$

$f_i \ge f_{i,\min}$, for all $i = 1, \cdots, n$

Kuhn-Tucker Condition

Non-linear Optimization Problem

Maximize $V(X), X \in R^n$

Subject to $C_j(X) \ge 0, 1 \le j \le k$

Kuhn-Tucker Conditions

$$\nabla\left(V - \sum_{j=1}^{k} \lambda_j C_j\right)(X) = 0$$

$$C_j(X) \ge 0, \lambda_j \ge 0$$

$$\lambda_j C_j(X) = 0$$

Proof

X^* is a maximizer of $\left(V - \sum_{j=1}^{k} \lambda_j C_j\right)(X)$

X^* satisfies all constraints $C_j(X) \ge 0$

For the X^*, $V(X) = \left(V - \sum_{j=1}^{k} \lambda_j C_j\right)(X)$

X* that satisfies the Kuhn-Tucker conditions is the optimal solution

Kuhn-Tucker Theorem

Non-linear Optimization Problem

Maximize $V(X), X \in R^n$

Subject to $C_j(X) \geq 0, 1 \leq j \leq k$

Kuhn-Tucker Conditions

$$\nabla\left(V - \sum_{j=1}^{k} \lambda_j C_j\right)(X) = 0$$

$$C_j(X) \geq 0, \lambda_j \geq 0$$

$$\lambda_j C_j(X) = 0$$

If both the objective function and the constraint functions are convex,

There exists a unique solution X* that satisfies the Kuhn-Tucker condition

Back to Our Problem ….

Maximize $\sum_{i=1}^{n} V_i(f_i) = \sum_{i=1}^{n} \omega_i \left(1 - e^{-\alpha_i f_i + \beta_i}\right)$

Subject to $\sum_{i=1}^{n} c_i f_i \leq 1$

$f_i \geq f_{i,\min}$, for all $i = 1, \cdots, n$

⇩

Kuhn-Tucker Conditions

$$\nabla\left(\sum_{i=1}^{n} \omega_i \left(1 - e^{-\alpha_i f_i + \beta_i}\right) - \lambda\left(1 - \sum_{i=1}^{n} c_i f_i\right) - \sum_{i=1}^{n} \lambda_i \left(f_i - f_{i,\min}\right)\right) = 0$$

$$1 - \sum_{i=1}^{n} c_i f_i \geq 0, \lambda \geq 0, \text{ and } f_i - f_{i,\min} \geq 0, \lambda_i \geq 0 \ (1 \leq i \leq n)$$

$$\lambda\left(1 - \sum_{i=1}^{n} c_i f_i\right) = 0 \text{ and } \lambda_i \left(f_i - f_{i,\min}\right) = 0 \ (1 \leq i \leq n)$$

Optimization (2)

$$f_i = \begin{cases} f_{i,\min} & i = 1,\cdots,p \\ \frac{1}{\alpha_i}\left(\beta_i + \ln\Gamma_i - Q\right) & i = p+1,\cdots,n \end{cases}$$

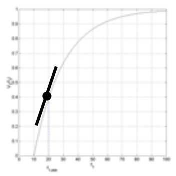
where $\Gamma_i = \frac{\omega_i \alpha_i}{c_i}, Q = \frac{1}{\sum_{i=p+1}^{n}\frac{c_i}{\alpha_i}}\left(\sum_{i=1}^{p} c_i f_{i,\min} + \sum_{i=p+1}^{n}\frac{c_i}{\alpha_i}\left(\beta_i + \ln\Gamma_i\right) - 1\right),$

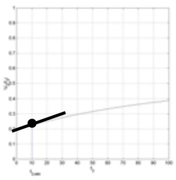
$f_1,\cdots,f_n$ are ordered according to $f_{1,\min},\cdots,f_{n,\min}$ which are arranged as

$$\Gamma_1 e^{-\alpha_1 f_{1,\min}+\beta_1} \le \Gamma_2 e^{-\alpha_2 f_{2,\min}+\beta_2} \le \cdots \le \Gamma_n e^{-\alpha_n f_{n,\min}+\beta_n},$$

and $p \in \{1,\cdots,n\}$ is the largest integer such that

$$\sum_{i=1}^{p} c_i f_{i,\min} + \sum_{i=p+1}^{n}\frac{c_i}{\alpha_i}\left(\alpha_p f_{p,\min} + \ln\frac{\Gamma_i}{\Gamma_p} + \beta_i - \beta_p\right) \ge UB.$$





Generalized Resource Management Framework (QRAM: Qos-based Resource Allocation Model)

- Quality depends on multiple factors (not only frequency)
  - Picture format (SQCIF, QCIF, CIF, 4CIF, 16CIF)
  - Color depth (1, 3, 8, 16, 24)
  - Frame rate (1, 2, …, 30)
  - Audio sampling rate (8, 16, 24, 44)
  - Audio bit count (8, 16)
- Discrete options (not continuous)
- Different types of resources (not only CPU)
  - CPU
  - Memory
  - Channel bandwidth

General Problem Formulation

n tasks : $i = 1,2,\cdots,n$

m QoS dimensions : $j = 1,2,\cdots,m$

L resources : $k = 1,2,\cdots,L$

Maximize $\sum_{i=1}^{n} V_i(q_{i,1}, q_{i,2}, \cdots q_{i,m})$

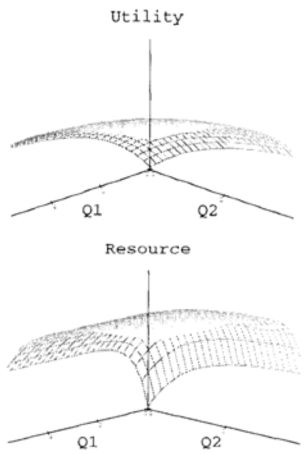
Subject to

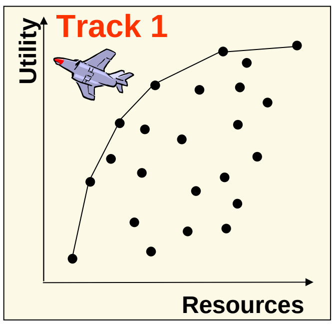
$\sum_{i=1}^{n} R_i^k(q_{i,1}, q_{i,2}, \cdots, q_{i,m}) \le RB^k$ for all resources $(1 \le k \le L)$

$q_{i,j} \ge q_{i,j,\min}$, for all tasks $(1 \le i \le n)$ and for all qos dimensions $(1 \le j \le m)$

Q-RAM Challenges

- QoS Option to V mapping
- QoS Option ro R mapping
- R to V mapping





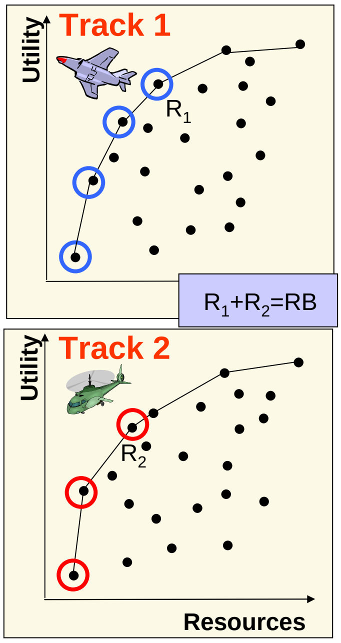
Optimization

Maximize $\sum_{i=1}^{n} V_i(R_i)$

Subject to

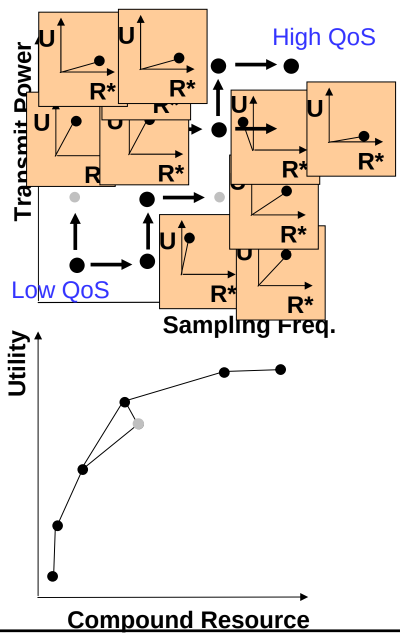
$$\sum_{i=1}^{n} R_i \le RB$$

$R_i \ge R_{i,\min}$, for all tasks $(1 \le i \le n)$

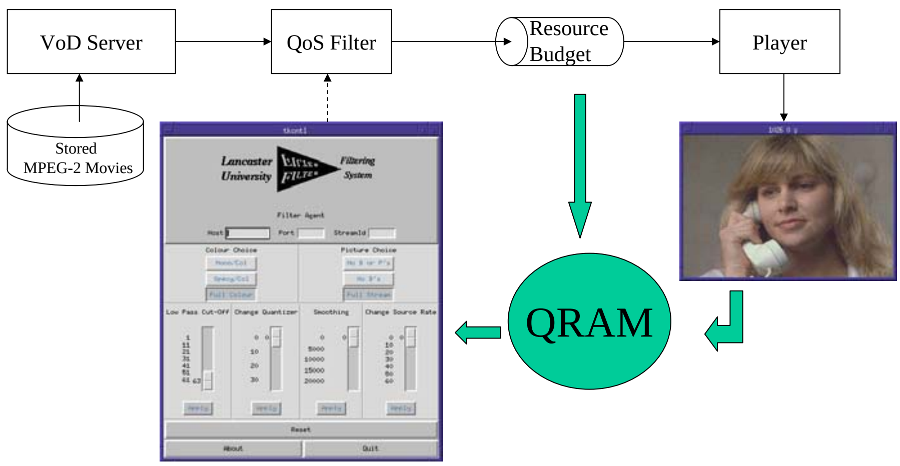


Online QoS Management

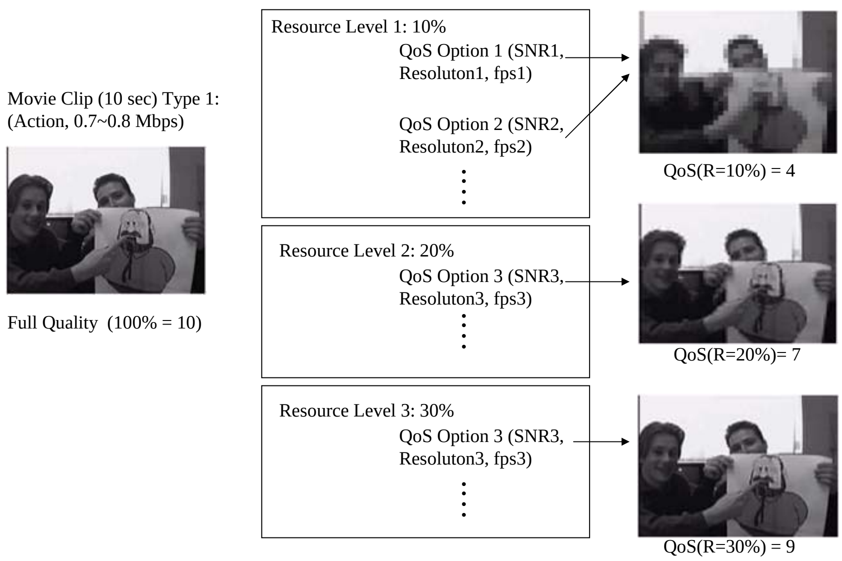
- Is it practial to use this optimization process for dynamically arriving/departing tasks?
  - Too complex to be applied online
- Quick composition of R-V relation



How to Apply to Multimedia



Depending on the resource availability….



References

- *Trade-off Analysis of Real-Time Control Performance and Schedulability* by Danbing Seto, John P. Lehoczky, Lui Sha, and Kang G. Shin, Real-Time systems, Vol. 21, Issue 3, November 2001
- *A Resource Allocation Model for QoS Management* by R. Rajkumar, C. Lee, J. Lehoczky, and D. Siewiorek, RTSS 1997.
- *Practical Solutions for QoS-based Resource Allocation Problems* by R. Rajkumar, C. Lee, J. P. Lehoczky, and D. P. Siewiorek, RTSS 1998