

# The Prefetch Aggressiveness Tradeoff in 360° Video Streaming

Mathias Almquist

Viktor Almquist

Vengatanathan Krishnamoorthi

Niklas Carlsson

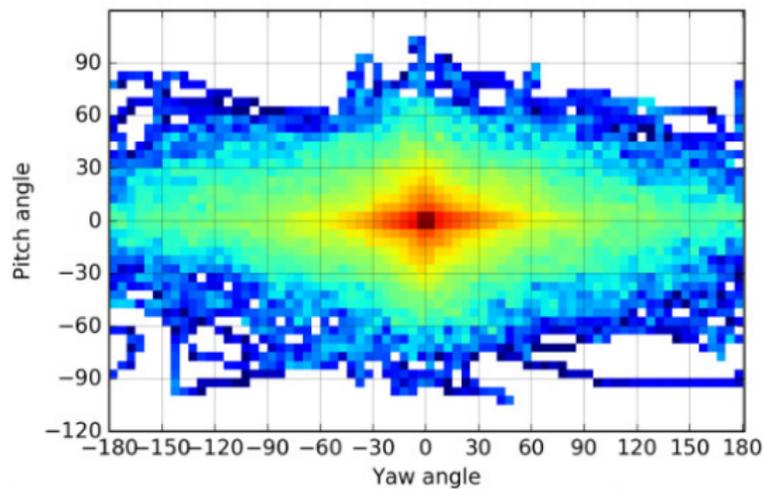
Derek Eager

# Intro

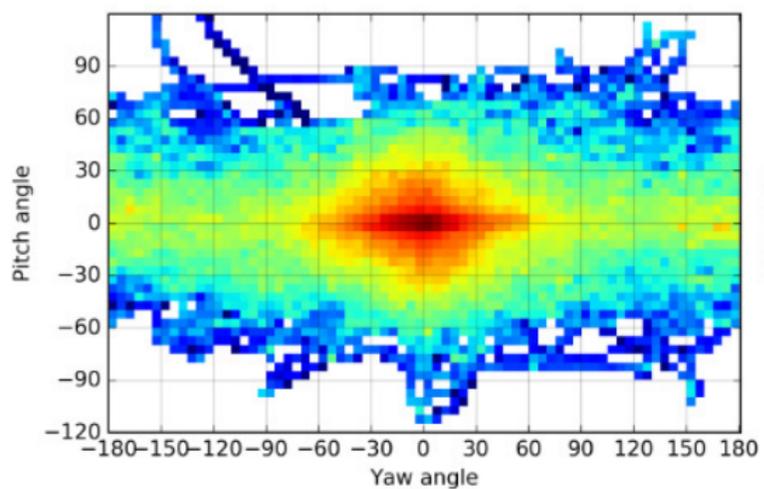
- With 360° video, only a limited fraction of the full view is displayed at each point in time.
- Delivering solutions that ignore the users' viewport waste bandwidth.
- Alternative playback quantities to be delivered for each candidate viewing direction.
- Video data must be buffered at the clients to protect playback stalls.
- Making good prefetching decision ↔ Protecting against future bandwidth variation

# Contribution

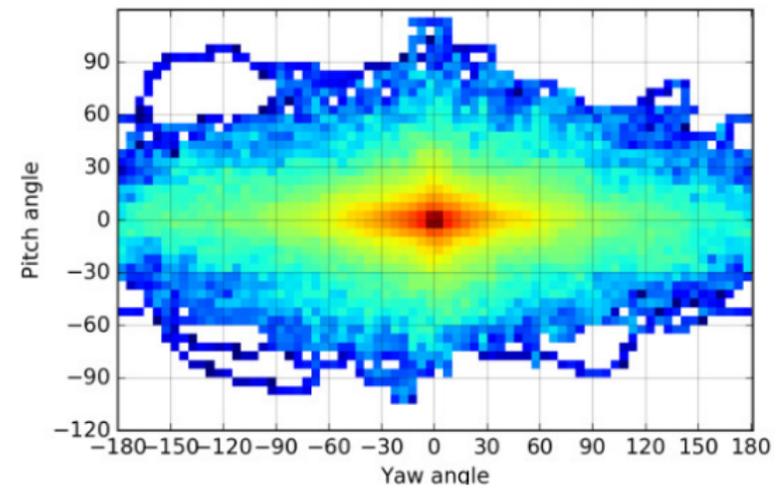
- An empirical characterization of head movement behavior based on data from viewing sessions of four different categories of 360° videos.
- An optimization-based comparison of prefetch aggressiveness tradeoffs seen for these video categories.



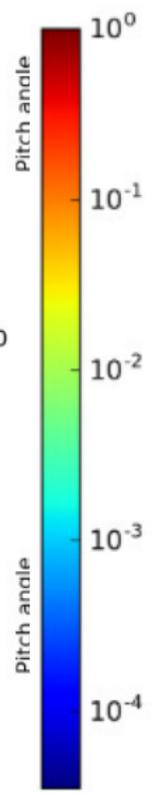
(a)  $T=2s$



(b)  $T=20s$



(c)

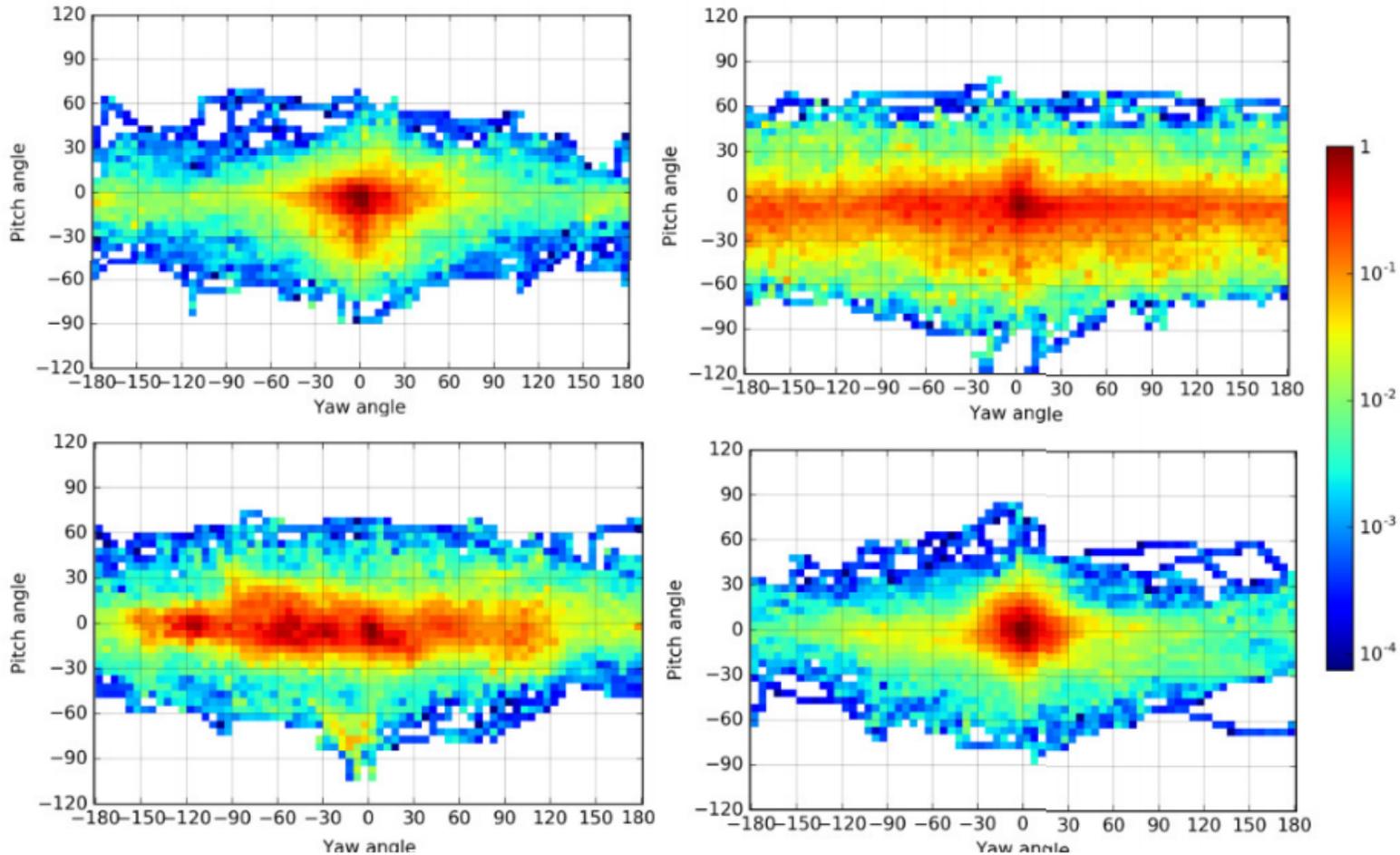


Example heat-maps for two categories of movies

# Measurement Methodology

- 30 4K 360° videos from Youtube (5 Categories: Exploration, Static focus, Moving focus, Rides, Miscellaneous)
- HMD: Oculus Rift CV1
- Recruit 32 people to watch the videos and record the orientations.

# Angular Utilizations-1

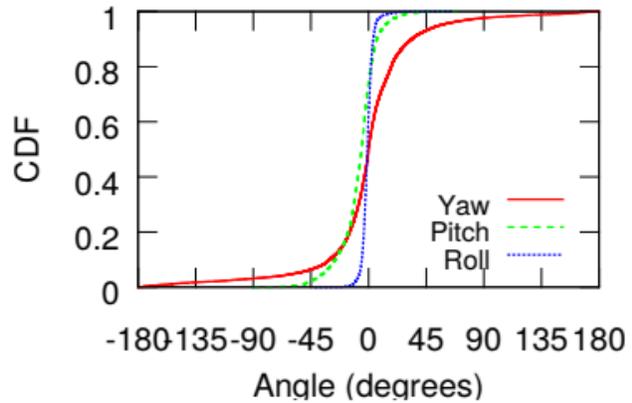


Rides( top-left)  
Exploration (top-right)  
Moving focus (bottom-left)  
Static focus (bottom-right)

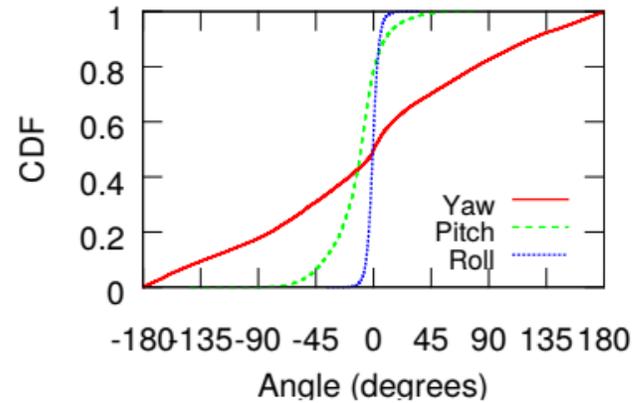
Only small changes over the  
video playback duration  
( 98% within  $10^\circ$ )

Heatmap of most utilized yaw and pitch angles

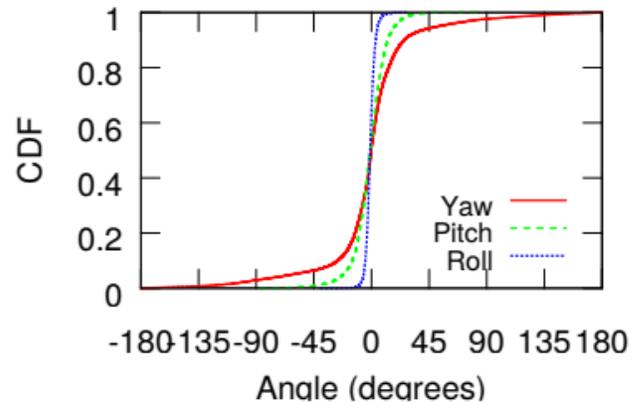
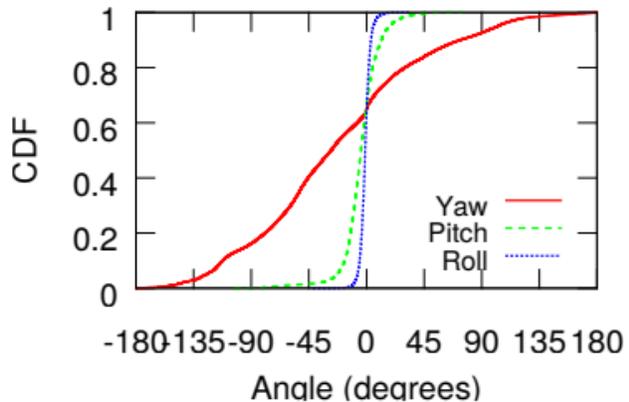
# Angular Utilizations-2



(a) Rides



(b) Exploration

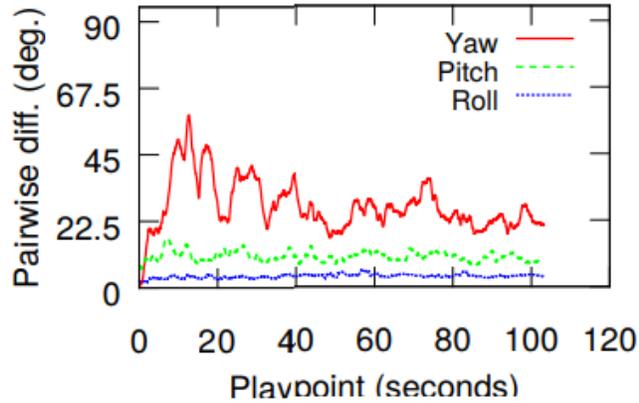


**Yaw is the most dominant orientation movement.**  
**⇒ Predicting the changes in yaw is the most important.**

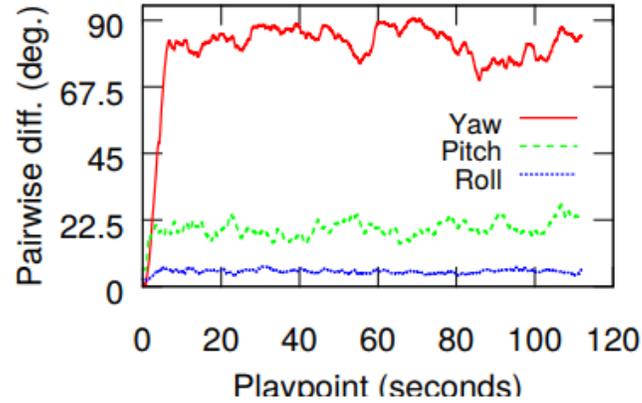
**Significant differences in the angular utilizations between the different categories.**

CDFs of angle utilization.

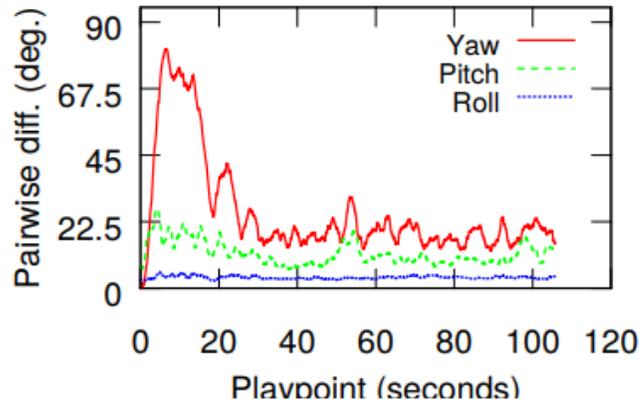
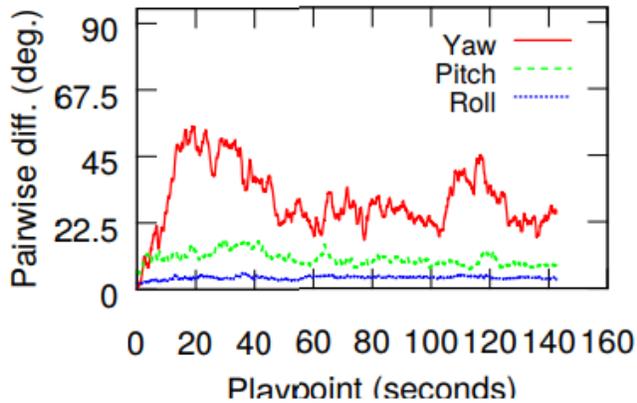
# Angular Utilizations-3



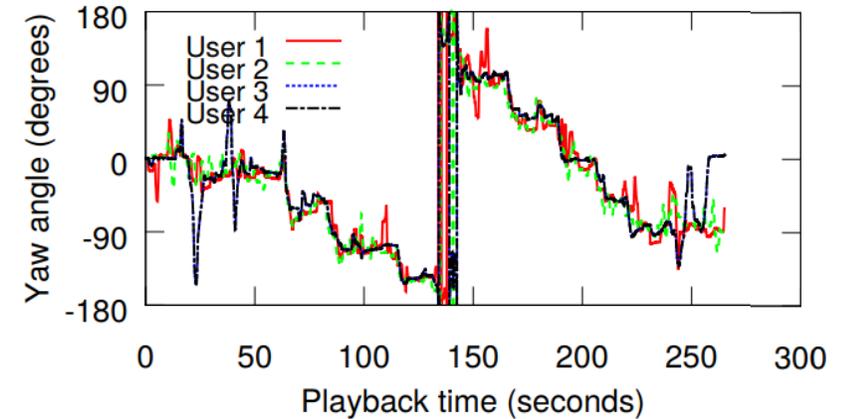
(a) Rides



(b) Exploration



Viewing patterns of other users are valuable when predicting the viewing direction of future viewers for some of the videos.



Average pairwise angular difference between users.



# Optimized Prefetching Tradeoffs

Symbol	Definition
$L$	Number of non-zero quality levels of the video
$N$	Number of tiles or discrete viewing directions
$x_{n,l}$	Binary variable indicating that the client will prefetch tile $n$ of the chunk at quality level $l$
$q_{n,l}$	Playback encoding of tile $n$ of the chunk with quality level $l$
$b_{n,l} = b(q_{n,l})$	Size of tile $n$ of the chunk with quality level $l$
$u_{n,l} = u(q_{n,l})$	Playback utility of playing tile $n$ (direction $n$ ) of the chunk with quality level $l$
$\Delta$	Playback duration of the chunk
$D$	Download rate
$T$	Average maintained buffer size in seconds
$E[u T]$	Expected playback utility, conditioned on $T$
$p_n(T)$	Probability looking in direction $n$ , $T$ time later
$\beta$	Parameter used to weight factors ( $0 \leq \beta \leq 1$ )

$$E[u|T] = \sum_{n=0}^{N-1} p_n(T) u(n|q_0, q_1, \dots, q_{N-1}), \quad (1)$$

$$p_n(T) = \int_{\theta_n}^{\theta_{n+1}} p(\theta|T) d\theta, \quad (2)$$

$\theta_n$  is the yaw angle corresponding to direction  $n$ , using CDFs.

# Detailed Optimization Model

Symbol	Definition
$L$	Number of non-zero quality levels of the video
$N$	Number of tiles or discrete viewing directions
$x_{n,l}$	Binary variable indicating that the client will prefetch tile $n$ of the chunk at quality level $l$
$q_{n,l}$	Playback encoding of tile $n$ of the chunk with quality level $l$
$b_{n,l} = b(q_{n,l})$	Size of tile $n$ of the chunk with quality level $l$
$u_{n,l} = u(q_{n,l})$	Playback utility of playing tile $n$ (direction $n$ ) of the chunk with quality level $l$
$\Delta$	Playback duration of the chunk
$D$	Download rate
$T$	Average maintained buffer size in seconds
$E[u T]$	Expected playback utility, conditioned on $T$
$p_n(T)$	Probability looking in direction $n$ , $T$ time later
$\beta$	Parameter used to weight factors ( $0 \leq \beta \leq 1$ )

$$\text{maximize } E[u|T], \quad (4)$$

where

$$E[u|T] = (1 - \beta) \left( \sum_{n=0}^{N-1} p_n(T) \sum_{l=0}^L x_{n,l} u_{n,l} \right) - \beta \left( \sum_{n=0}^{N-1} \frac{p_n(T) + p_{n+1}(T)}{2} \right. \\ \left. \times \sum_{l=0}^L \sum_{l'=0}^L x_{n,l} x_{n+1,l'} |u_{n,l} - u_{n+1,l'}| \right), \quad (5)$$

such that

$$\sum_{l=0}^L x_{n,l} = 1, \quad 0 \leq n < N, \quad (6)$$

$$\sum_{n=0}^{N-1} \sum_{l=0}^L x_{n,l} b_{n,l} \leq D\Delta, \quad (7)$$

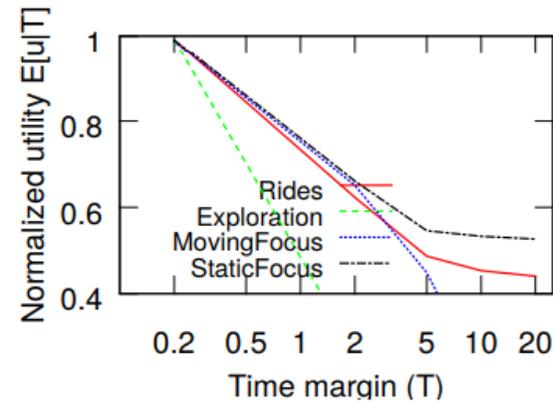
$$x_{n,l} \in \{0, 1\}, \quad 0 \leq n < N, 0 \leq l \leq L. \quad (8)$$

$$u(q) = b \cdot \frac{(q/\theta)^{1-a} - 1}{1-a}$$

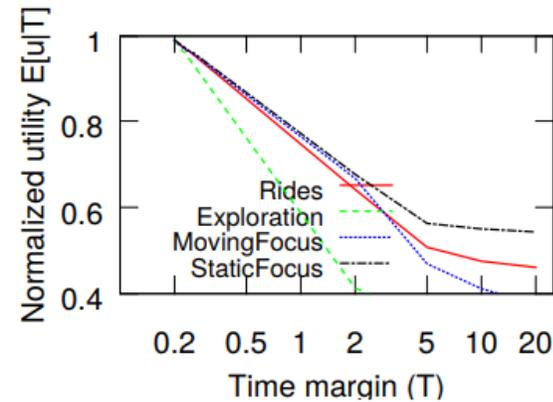
$$u_{n,0} = -f u_{n,L}:$$

where  $f$  is a penalty factor.

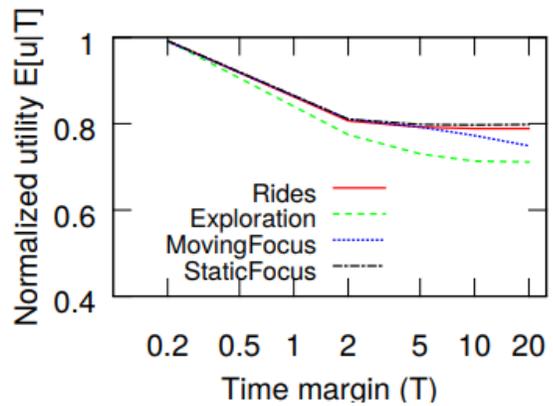
# Example Characterization-1



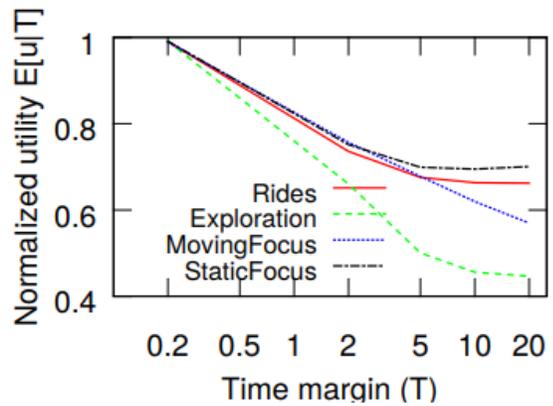
(a) Linear utility



(b) Square-root utility



(c) Logarithmic utility

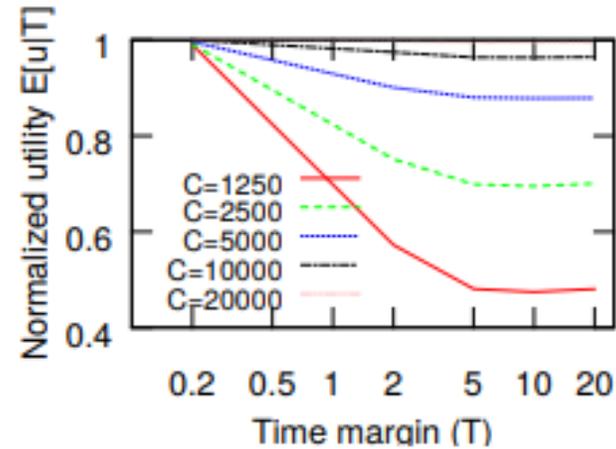


(d) Large-screen utility

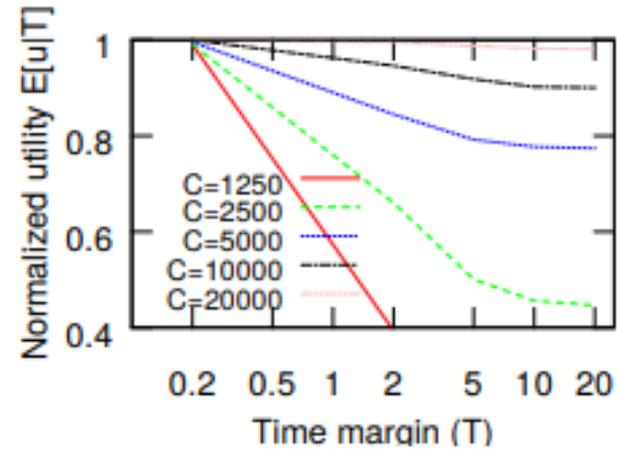
Example tradeoffs for different 360° video categories.  
( $C = 2500$  in all sub-figures.)  
( $C$ : the number of bytes that can be downloaded during a time slot)

**Moving focus, Exploration => short-term prefetching**  
**Rides, Static Focus => long-term prefetching**

# Example Characterization-2



(a) Static focus



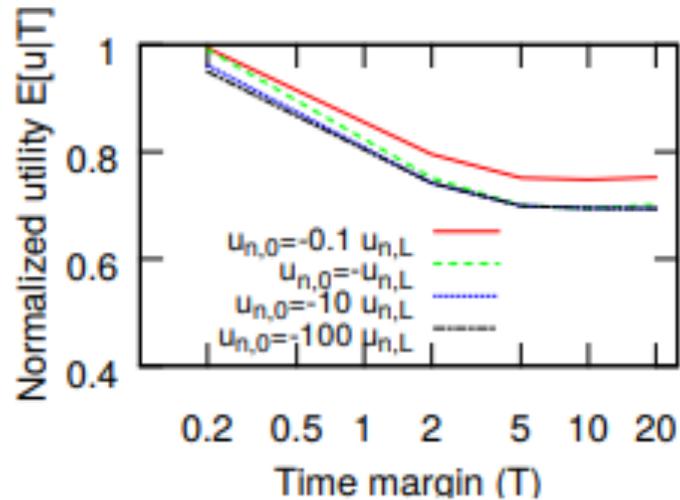
(b) Exploration

**Figure 11: Example tradeoffs for different capacities. (Large-screen utility function with  $\frac{u_{n,0}}{u_{n,1}} = -1$  and  $\beta = 0.001$ .)**

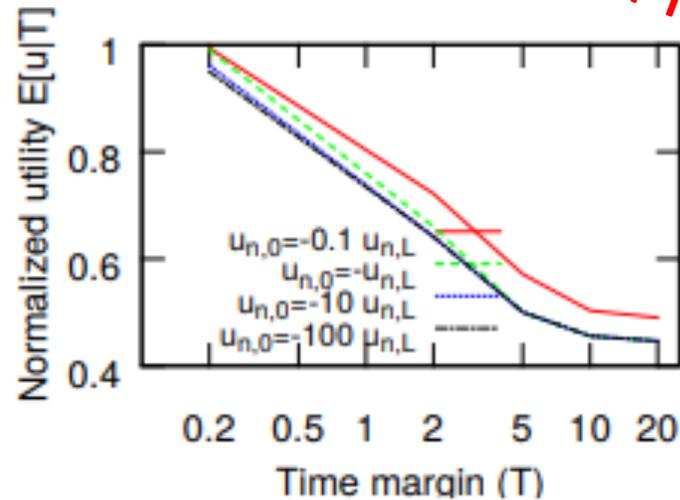
Diminishing prefetch capacity returns

# Example Characterization-3

Limited impact of stall penalty



(a) Static focus



(b) Exploration

**Figure 12: Example tradeoffs for different penalties  $\frac{u_{n,0}}{u_{n,L}}$ . (Large-screen utility function with  $C = 2500$  and  $\beta = 0.001$ .)**