QAVA: Quota Aware Video Adaptation

Jiasi Chen, Amitabha Ghosh, Josphat Magutt, Mung Chiang

Princeton University

Dec. 12, 2012

Rise of Usage-Based Pricing

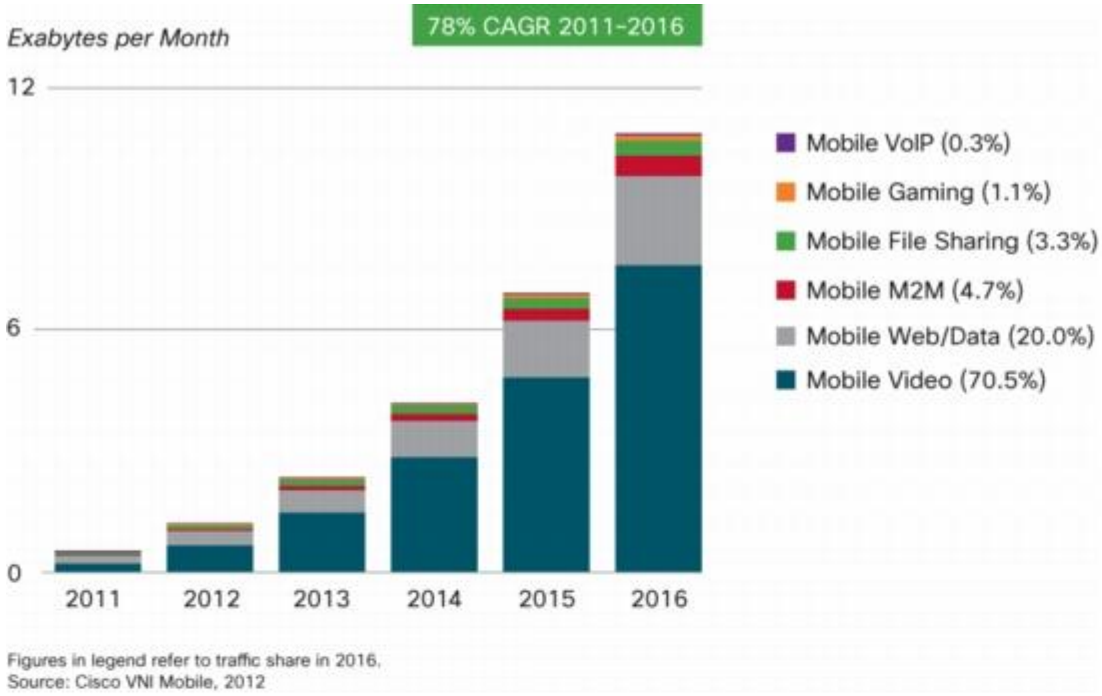
10

$/GB charged by AT&T Wireless
for 3G/4G data usage above 2GB

Rise of Video Traffic

70

Percentage of mobile data from video in 2016



Source: Cisco Visual Networking Index 2012

The Conflict Between Two Trends

Two emerging trends of Internet application:

- Video traffic becoming dominant

  High-resolution devices (*e.g.*, iPhone, iPad, Android tablets)

	Upstream Traffic		Downstream Traffic		Total Traffic	
Rank	Application	Share	Application	Share	Application	Share
1	BitTorrent	52.01%	Netflix	29.70%	Netflix	24.71%
2	HTTP	8.31%	HTTP	18.36%	BitTorrent	17.23%
3	Skype	3.81%	YouTube	11.04%	HTTP	17.18%
4	Netflix	3.59%	BitTorrent	10.37%	YouTube	9.85%
5	PPStream	2.92%	Flash Video	4.88%	Flash Video	3.62%

SOURCE: SANDVINE NETWORK DEMOGRAPHICS

sandvine

- Usage-based pricing becoming prevalent

Carrier	Country	Wireline/Wireless	Baseline Quota	Overage Charge
AT&T	USA	Wireless	2 GB	10 USD per GB
Verizon	USA	Wireless	2 GB	10 USD per GB
Reliance	India	Wireless	2 GB	0.01 Rupee per 10 kB
Rogers	Canada	Wireline	80 GB	2 CAD per GB
AT&T	USA	Wireline	250 GB	10 USD per 50 GB

Can the consumer consume content without worrying about her wallet?

Current Video Adaptation Solutions

Two main approaches:

- Consumers may be warned by service providers or applications
  - Android 4.0 provides data usage monitoring app; other iOS / Android apps
- "One size fits all" cutting back bit rates across all videos, for all users, at all times
  - Youtube: channel-based quality adaptation depending on connection type
  - Netflix: static quality adaptation to address wireline ISP quota constraints

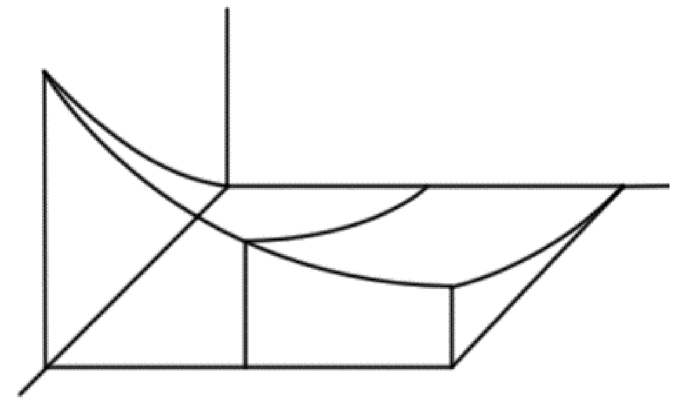
Onavo: mobile app that compresses images and text to use less data

Adaptive HTTP streaming for bandwidth constraints

- Adobe Dynamic Streaming for Flash
- Microsoft Smooth Streaming for Silverlight and Windows Phone
- Apple HTTP Live Streaming for iOS

Video Consumption Tradeoff

A 3-way tradeoff



Within budget	Minimize	Supply
Cost	Distortion	# Videos watched
↓	↓	↓
MB of video	Video compressibility	Usage profile

Quota-Aware Video Adaptation (QAVA)

Is every bit needed for every user at every time?

Key idea: All bytes are *charged* the same on cellular data plans, but not all bytes are equally *valuable* to mobile video experience.

Toy example: http://www.youtube.com/watch?v=0sUBDpS9e2U

Stream Selector

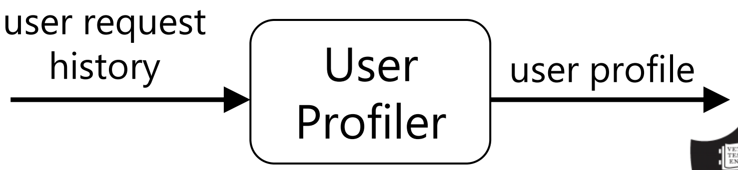
Choose the right bitrate to maximize video quality

Video Profiler

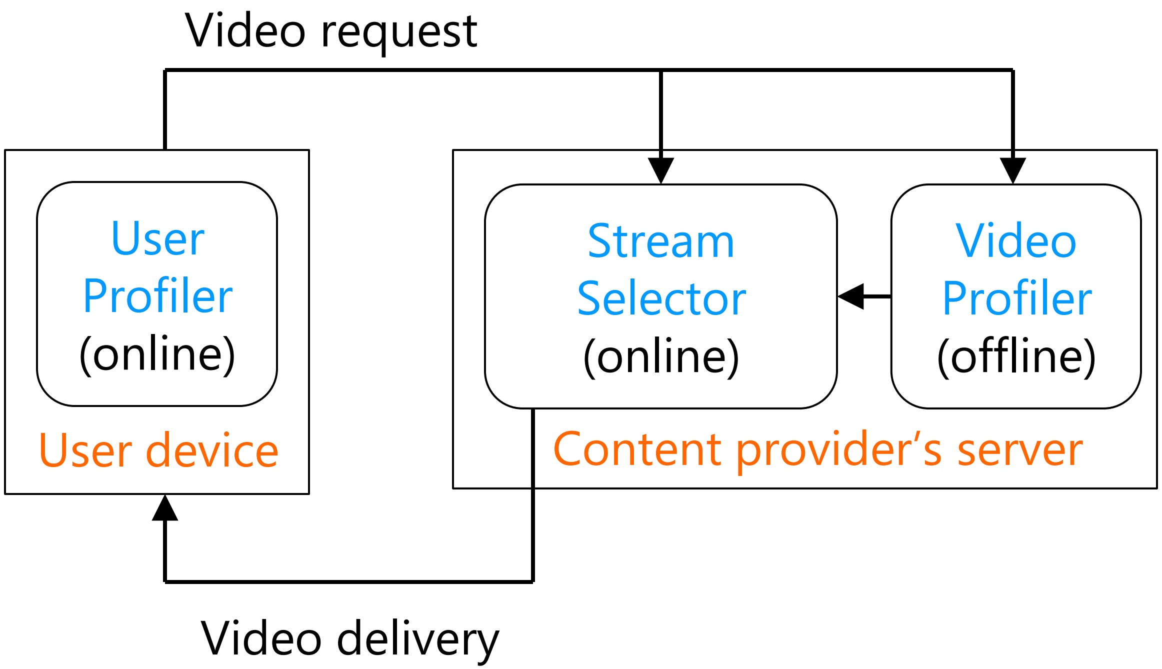
Estimate compressibility of video

User Profiler

Predict user's behavioral patterns from past history



QAVA System Architecture



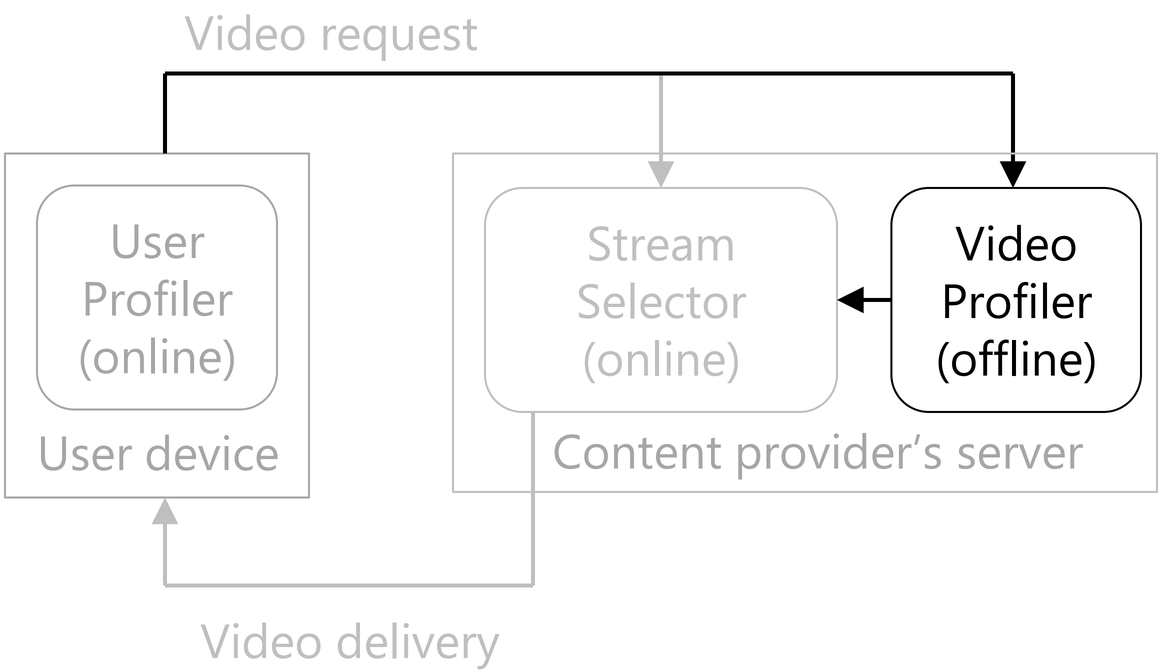
Stream selector: located on user device / network / content provider

User profiler: requires access to user request logs

Video profiler: requires access to videos

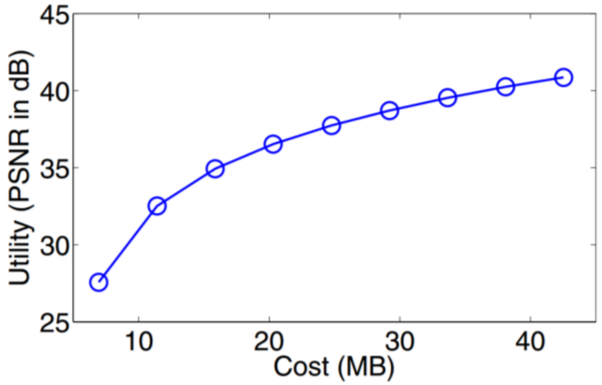
Video Profiler

Estimate video compressibility



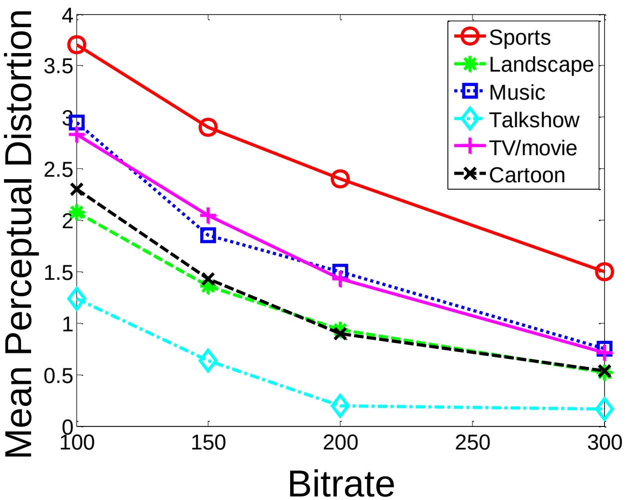
Leveraging Video Compressibility

Utility-cost tradeoff: diminishing returns for increasing cost



H.264/AVC video
Encoded at 100-900 kbps
720×480 pixels
Duration 6 mins

Different types of videos have different tradeoff curves – leverage this!



H.264/AVC videos
Encoded at 100,150,200, 300 kbps
640x480 pixels

Video Compressibility Demo

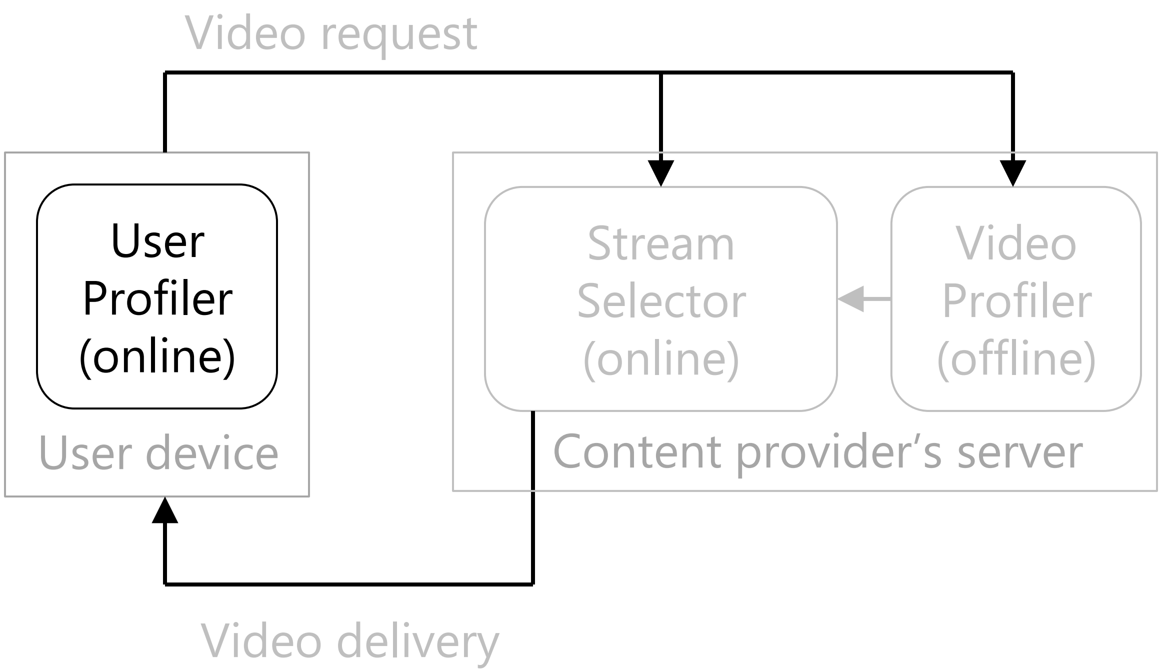
http://www.youtube.com/watch?v=bnGGS_u5doo

1. Talk show
   Left: 300 kbps
   Right: 100 kbps
2. Action clip
   Left: 100 kbps
   Right: 300 kbps

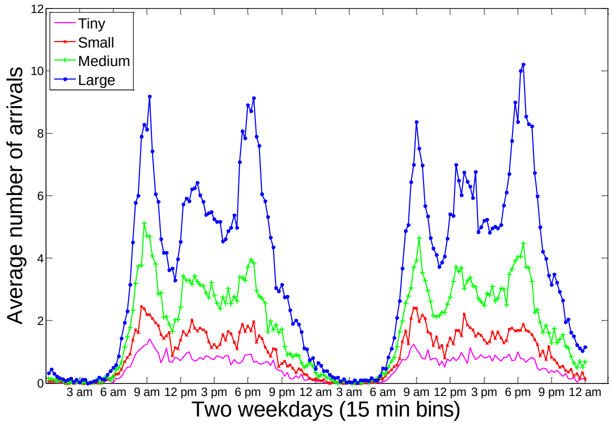
Takeaway: Users have different perception of low- and high-motion videos. Low-motion videos are more compressible without perceptually noticeable distortion.

User Profiler

Predict user's future data consumption patterns



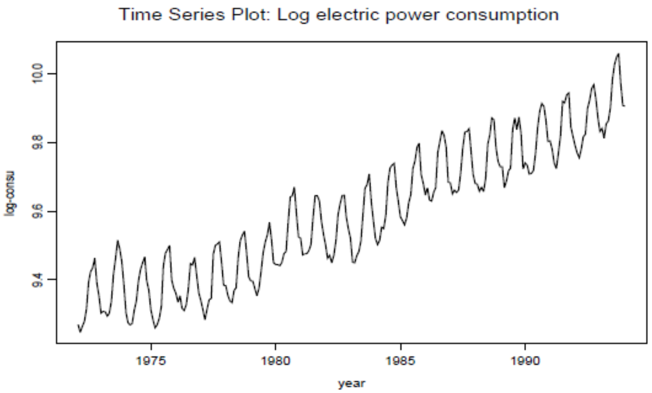
Seasonality and Trend in Time Series



Seasonality

Regularly spaced peaks and troughs with a consistent direction and approximately the same magnitude

Customer arrival in Starbucks who use Wi-Fi, NYC March 2010



Trend

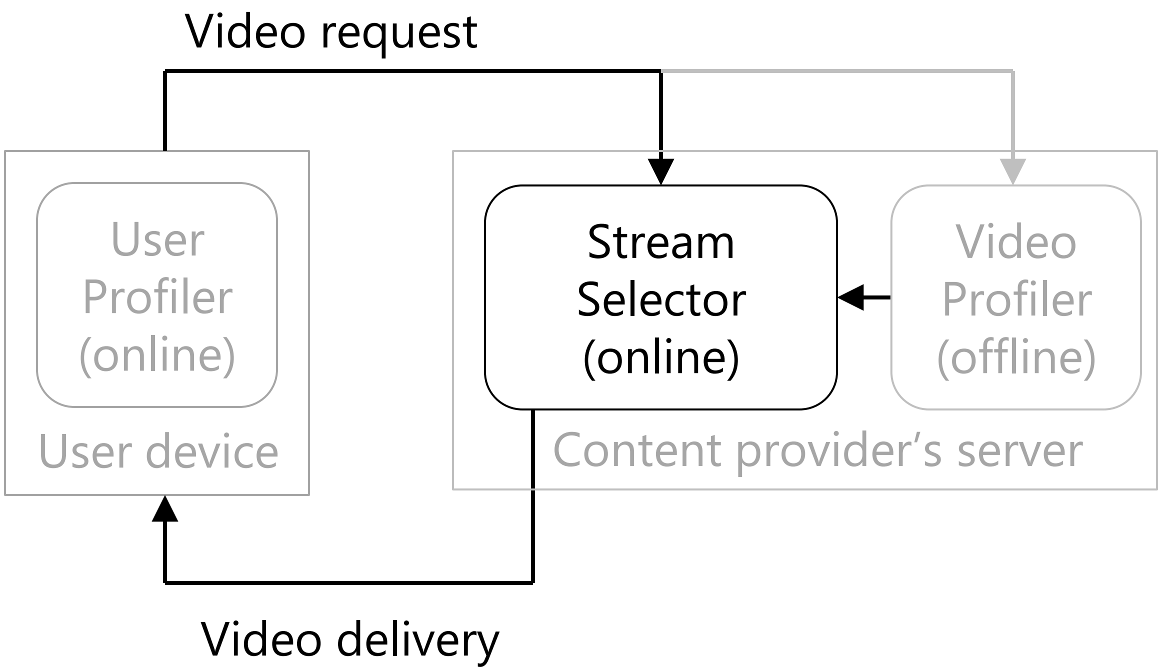
Long term movement with an underlying upward or downward direction

Electric power consumption between 1975 and 1990

Our approach: estimate probability of request arrival in each time period
estimate video type preferences of each user

Stream Selection

How to choose the delivered video bitrate while staying under quota?



Offline Stream Selection

If all video requests are known, we have the offline problem:

$$\underset{x_{tj}}{\text{maximize}} \quad \sum_{t=1}^{T}\sum_{j=1}^{M_t} u_{tj}x_{tj}$$

maximize the total / average utility

$$\text{subject to} \quad \sum_{t=1}^{N}\sum_{j=1}^{M_t} c_{tj}x_{tj} \le B$$

spend less than budget

$$\sum_{j=1}^{M} x_{tj} \le 1, \ \forall\, t$$

choose at most one bitrate per video

$$x_{tj} \in \{0,1\}, \ \forall\, t, j$$

B: quota budget
T: number of time periods
M_t: # of versions of video t
u_{tj}: utility of version j of video t
c_{tj}: cost of version j of videot
x_{tj}: 1 if version j of video t is selected; 0 otherwise

This is the multiple-choice knapsack problem

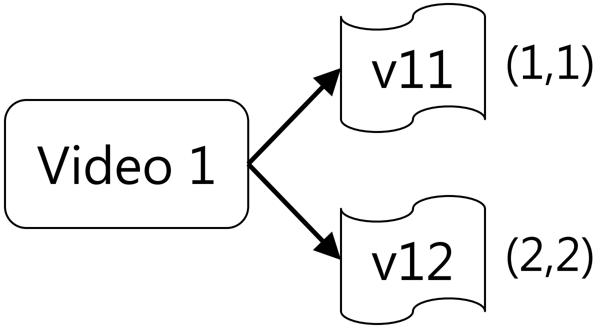
Kellerer H, Pferschy U, Pisinger D, *Knapsack Problems*, Springer 2004

Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)

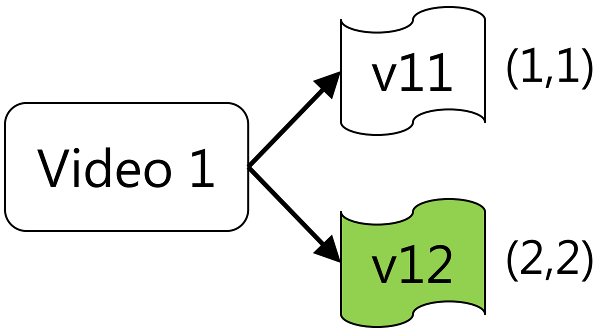


Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)

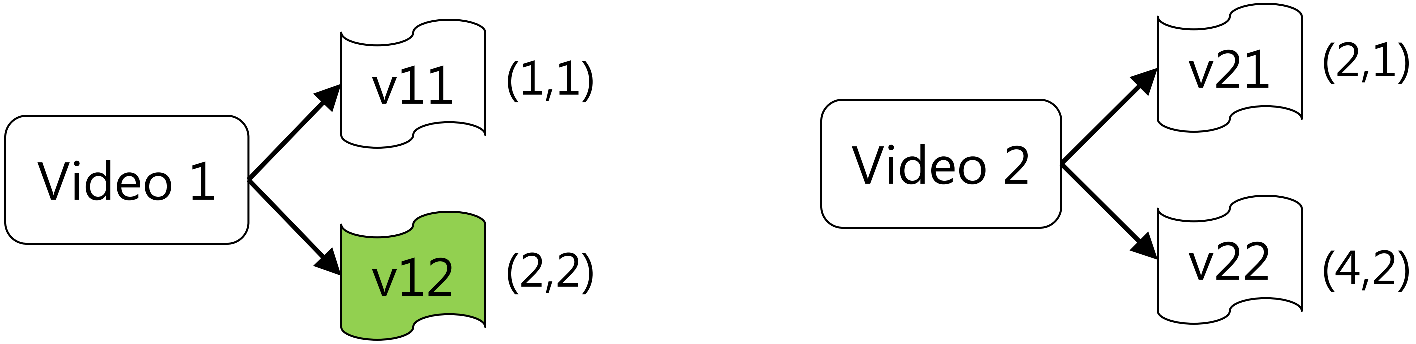


Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)

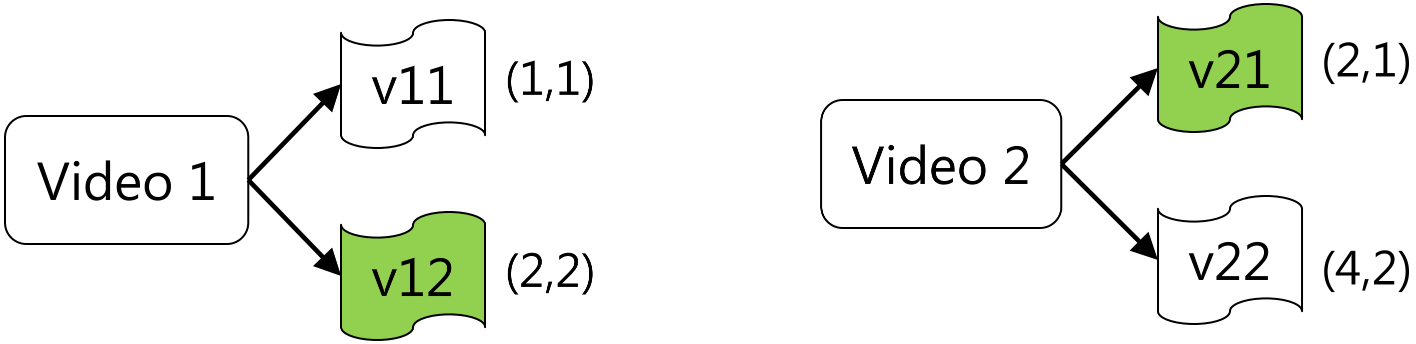


Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)

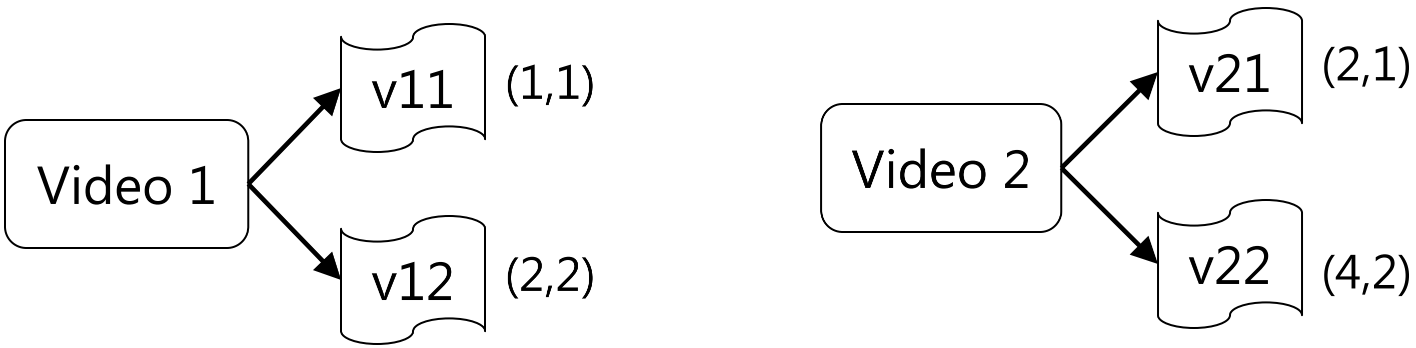


Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)

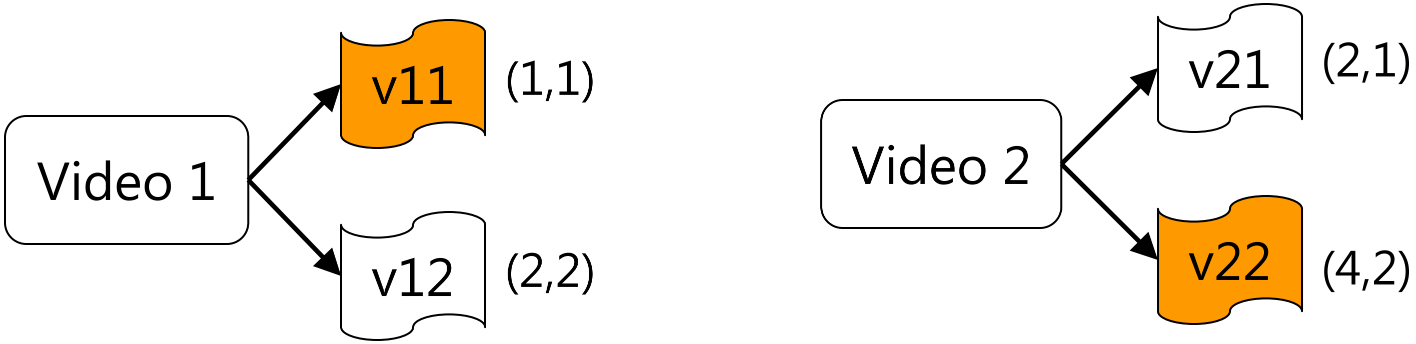


Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)

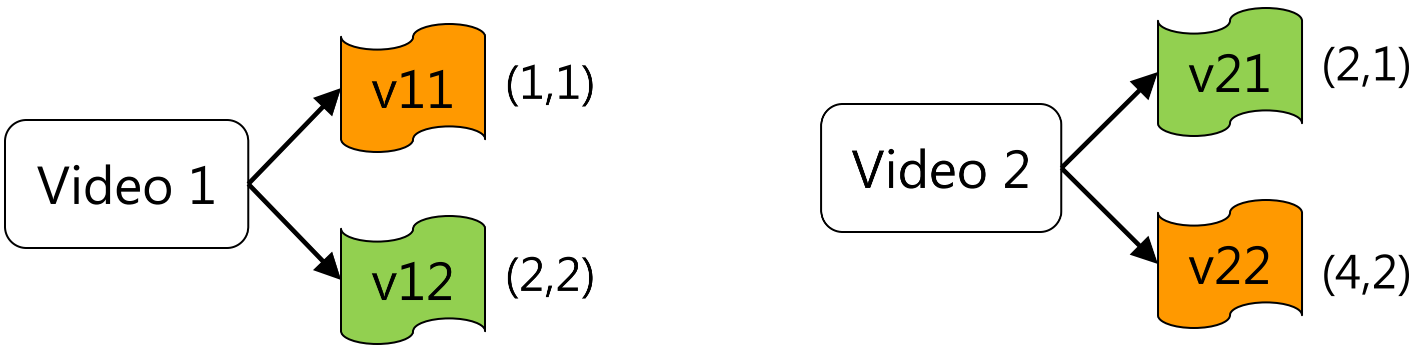


Online vs. Offline Stream Selection

Budget: 3

Goal: Maximize total utility

Items: (utility, cost)



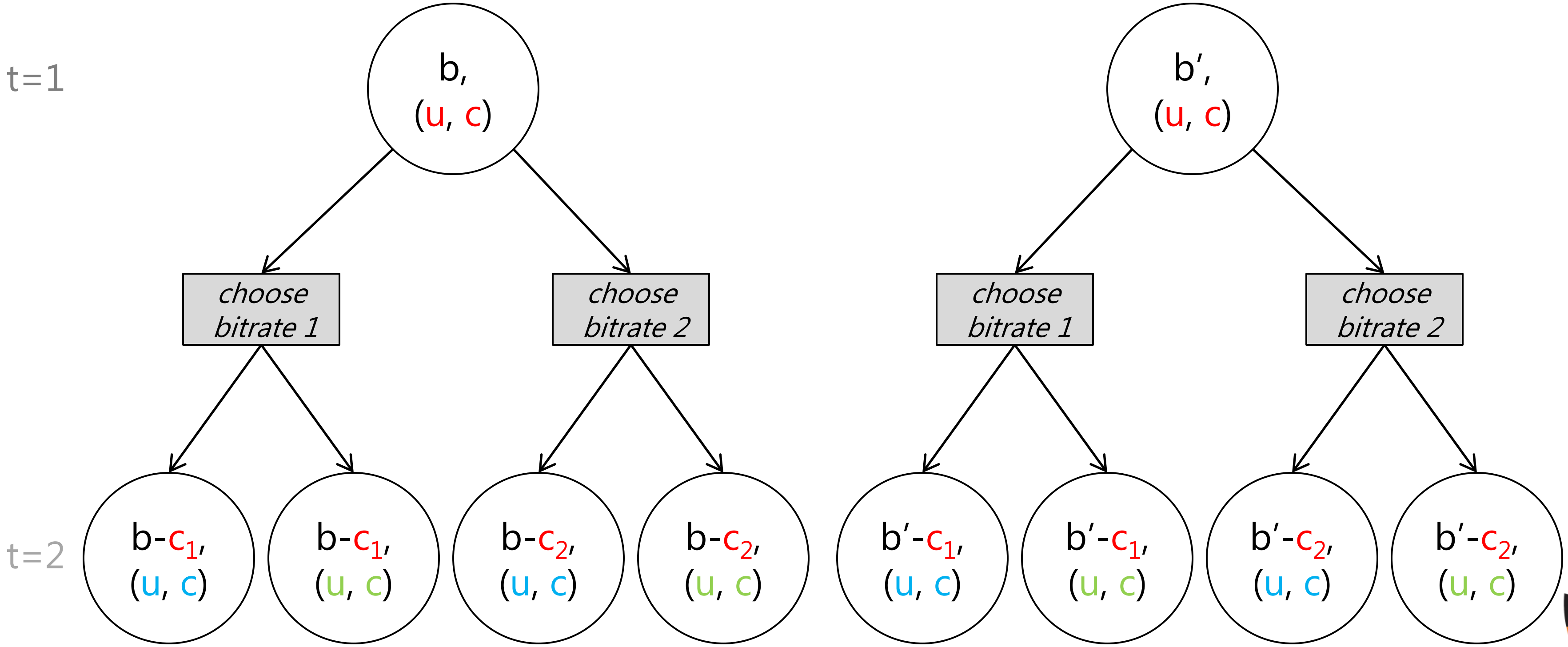
Offline optimal: v11, v22
Total utility: 1+4 = 5
Total cost: 1+2 = 3

Online greedy: v12, v21
Total utility: 2+2 = 4
Total cost: 2+1 = 3

Modeling using Markov Decision Process

Possible videos V = { (u,c), (u,c), (u,c) }; videos arrive randomly
Which bitrate to choose?

Markov property: Future bitrate decisions depend only on remaining budget, independent of past bitrate decisions



Simulation using Video Request Traces

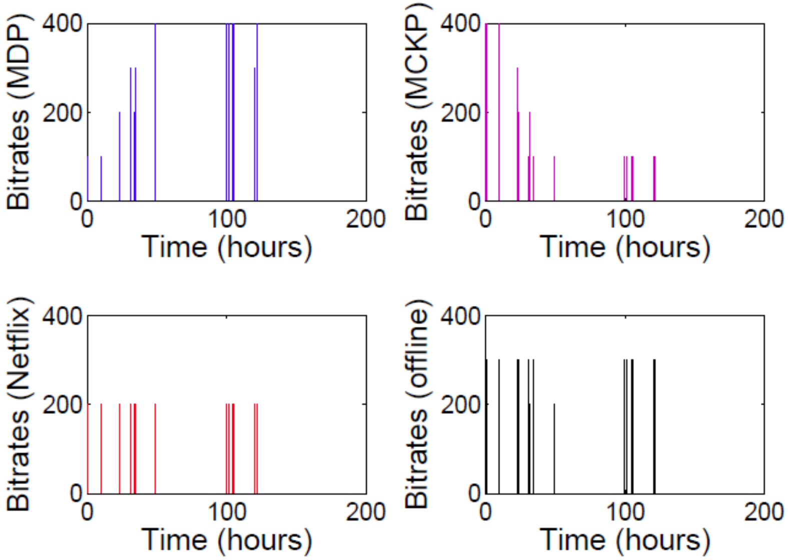
YouTube request traces from wireless campus network

- 14 days, 16 337 users, 611 968 requests

4 bitrate selection algorithms:

- MDP: Our proposed approach
- MCKP: State-of-the art literature
- Netflix: Solution in practice

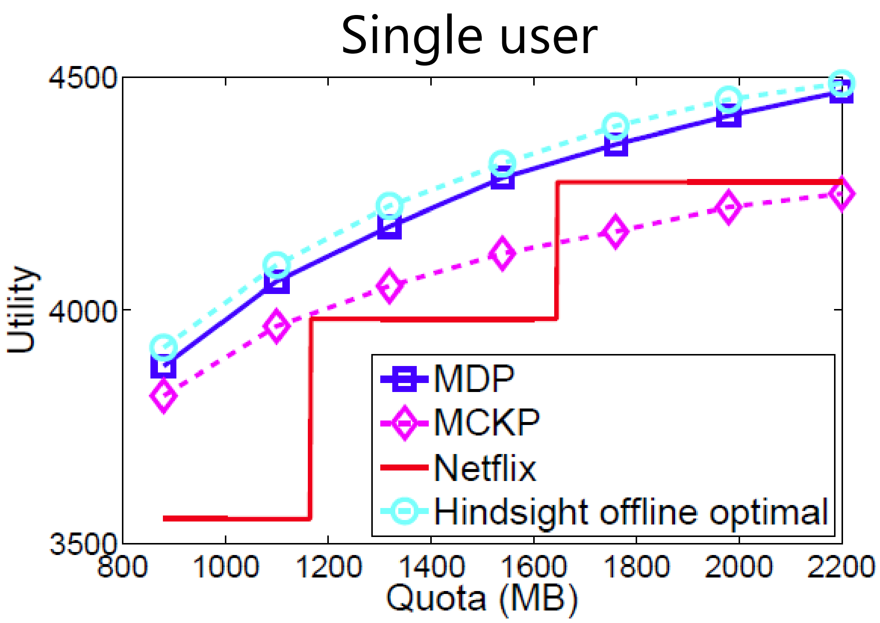
  Caveat: assumes perfect knowledge of number of video requests
- Offline: Hindsight offline optimal

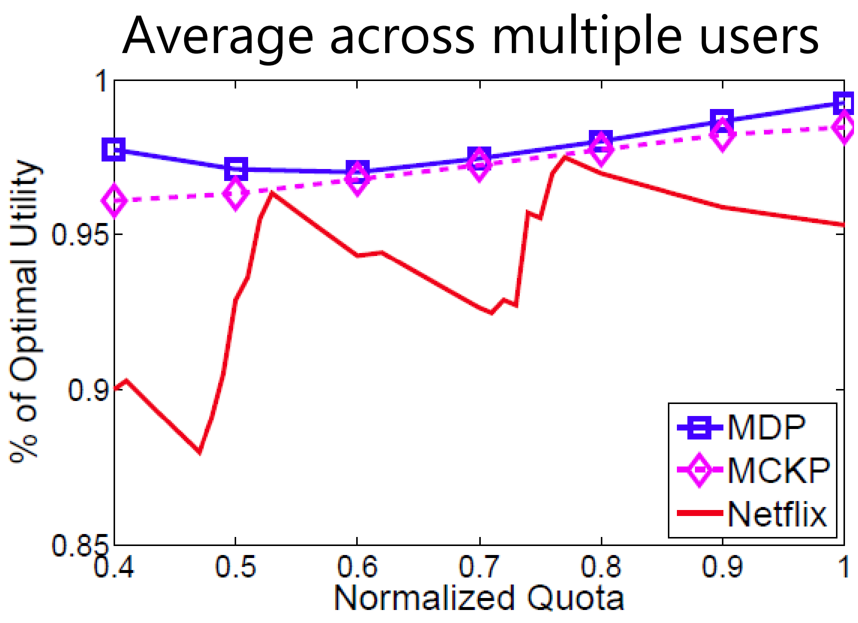


Zink M, Suh K, Gu Y, Kurose J, "Watch Global Cache Local: YouTube Network Traces at a Campus Network - Measurements and Implications", *IEEE MMCN*, 2008.

Stream Selection Algorithm Comparison

How do algorithms perform for different user request traces, sweeping across quotas?

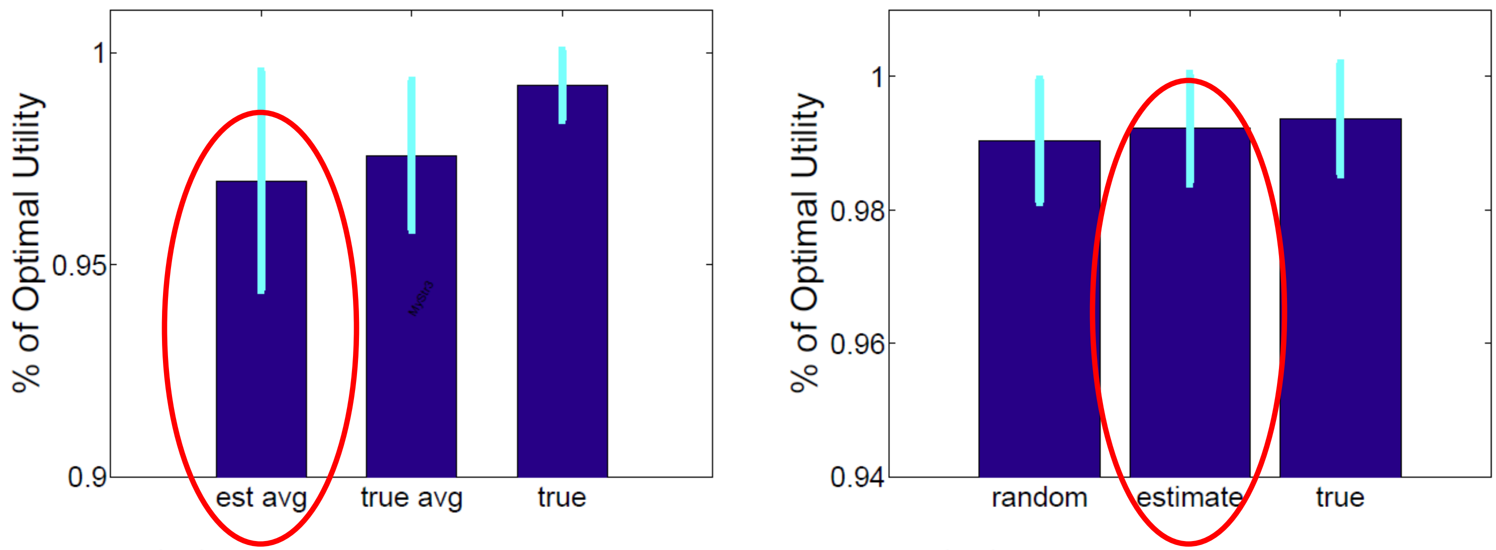




Conclusion: MDP achieves greater utility than other algorithms, without exceeding the quota

Effects of Prediction Error

How robust is MDP algorithm to wrong user profiler or video profiler information?



(a) Arrival rate (b) Video type

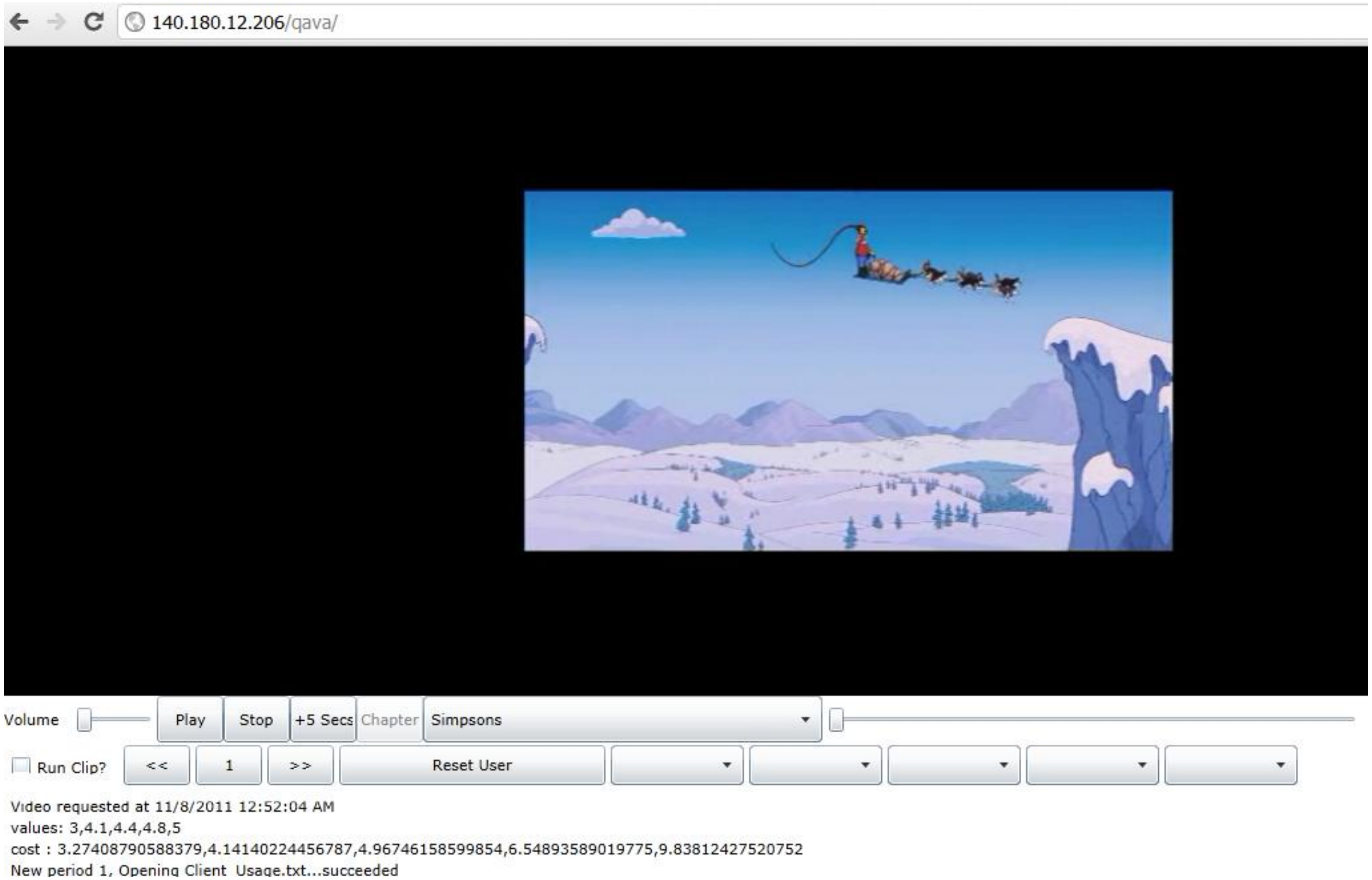
<u>Conclusion</u>: Incorrect information only slightly decreases solution optimality

Implementation

Goals

- Test our architecture and system design
- Understand consumption behavior of real people
- Understand user perception of video quality
- Evaluate the algorithm
- Fun to run a trial involving real people

Silverlight Web Browser



Proof-of-concept implementation in web browser using Microsoft Silverlight

Android App Volunteer Trial

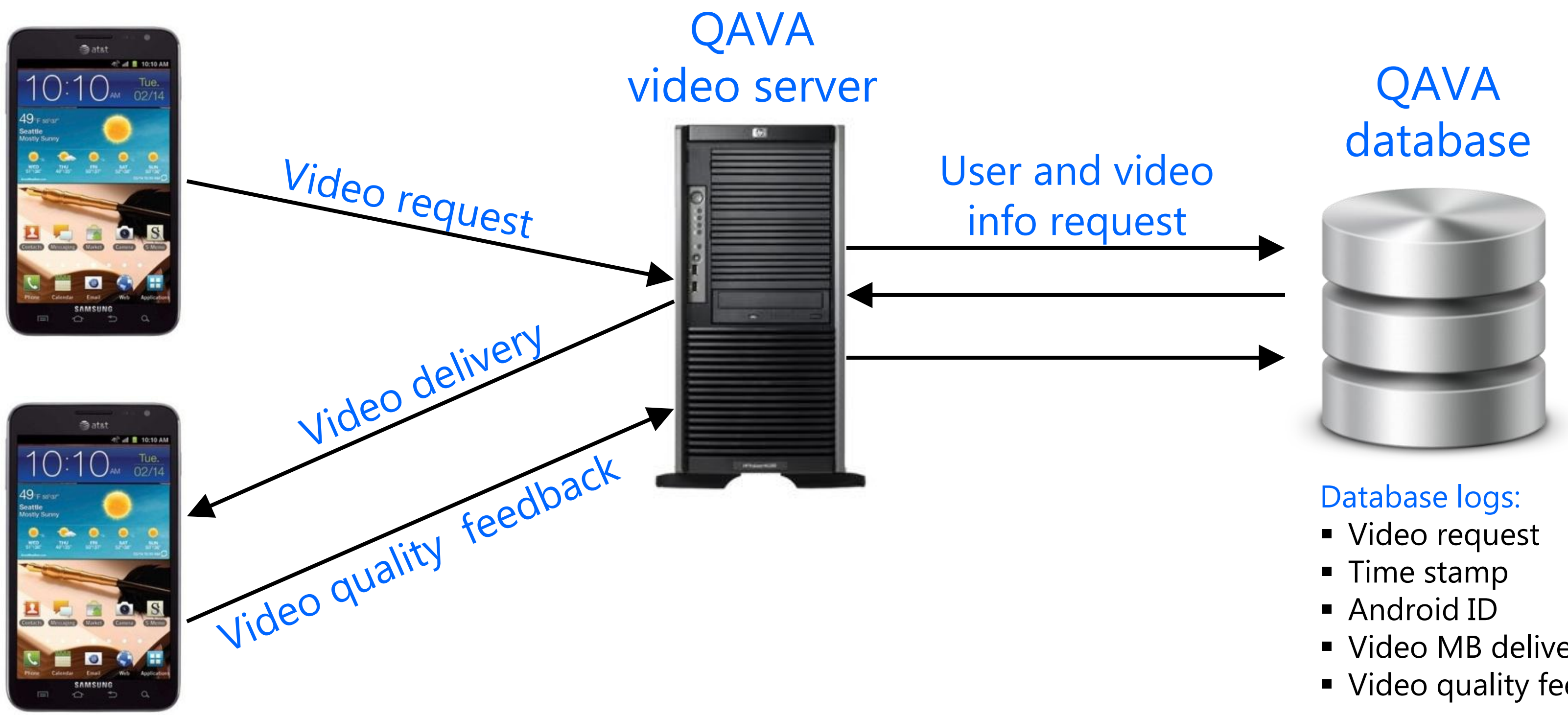
Developed QAVA as an Android application

Content provider: QAVA server

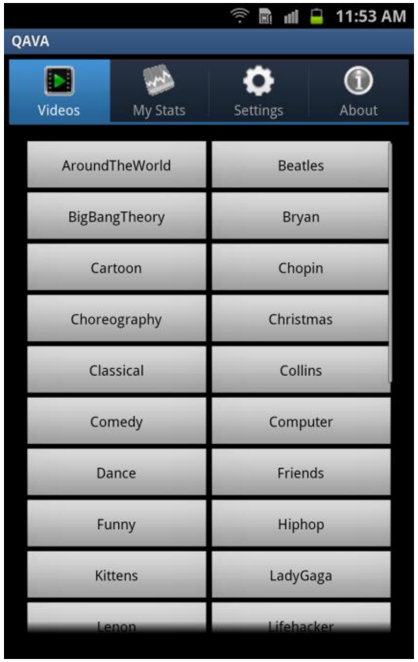
~500 videos encoded at 25 Kbps granularity (100 Kbps – 500 Kbps)

Participants: ~15 volunteers from Princeton community

Evaluation: Video quality feedback from users

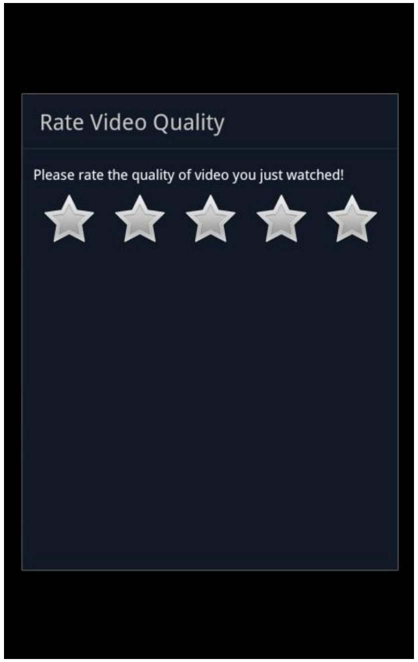


Android App Screenshots



Category selection
Tailored to user preferences

Video selection
Regularly updated with new content



Video feedback
Primary means of evaluating user satisfaction

Conclusions & Future Work

Discussed conflicting trends of:

- Usage based pricing
- Increasing video consumption

Developed system design for quota-aware video adaptation

- Key idea: Not every bit needed for every user at every time
- Compared state-of-the-art literature and practical algorithms for video rate adaptation

Next: evaluate system performance with real user trial

explore client-based implementation architectures

Thank you!

QUESTIONS?