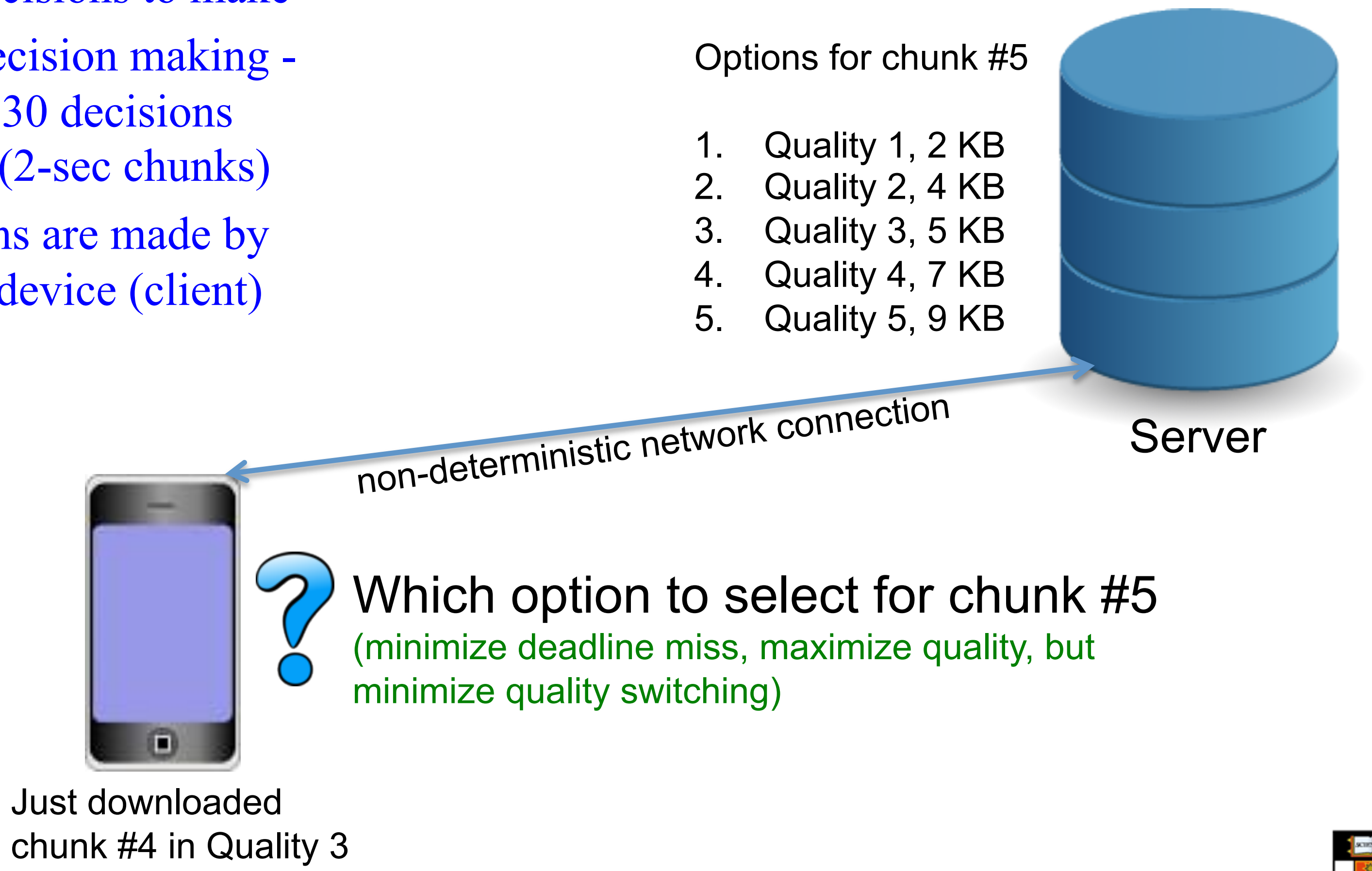
HTTP-based Adaptive Streaming for Mobile Clients using Markov Decision Process

Ayub Bokani Mahbub Hassan Salil Kanhere

Challenges of Decision Making in DASH

- Difficult decisions to make
- Frequent decision making - as much as 30 decisions per minute (2-sec chunks)
- All decisions are made by the mobile device (client)



Markov Decision Process

- A tool for optimizing decision making when outcomes are partly random and partly under the control of the decision maker
- The process goes through a finite set of states
- At each state, the decision maker can choose a particular action from a given set
- State transitions are random, but the transition probabilities are different for different actions
- Each action is associated with a reward and sometimes a penalty
- Revenue (award minus penalty) is used to evaluate the outcome for a given action taken at a given state
- Solving an MDP problem means finding the best action for each state that will maximize the overall revenue
- Can be solved using dynamic programming
- Once solved, the actions corresponding to the states is called the optimal strategy

DASHing with MDP

- Researchers have confirmed that use of MDP improves streaming quality
- However, MDP increases computation overhead of decision making at the mobile
- The overhead is an issue if a new MDP problem has to be solved for each chunk
- Our goal is to explore different approaches to reduce decision making overhead for MDP-based DASH

Presentation Overview

- Our MDP model for DASH
- 3 approaches to reduce MDP overhead
- Simulation setup for evaluating these approaches
- Results
- Conclusion

MDP Formulation of DASH (1)

- States S(ρ,q): q is quality level of the downloaded chunk and ρ is the time available before the chunk's playback deadline (*current buffer occupancy measured in time*)
- Actions (decisions): quality level of the next chunk to be downloaded
- Rewards: higher reward for watching a chunk in higher quality
- Penalties: there is a penalty for missing a deadline as well as switching quality from previous chunk to the next

quality level (q)	**1**	**2**	**3**	**4**	**5**
$u(q)$	1	2	4	7	10

Rewards for quality levels

Penalties for switching quality levels

$i \backslash j$	**1**	**2**	**3**	**4**	**5**
1	0	1	5	10	25
2	10	0	1	5	10
3	50	10	0	1	5
4	250	50	10	0	1
5	500	250	50	10	0

Penalty for missing a deadline = probability of deadline miss x D

MDP Formulation of DASH (2)

The role of Bandwidth CDF

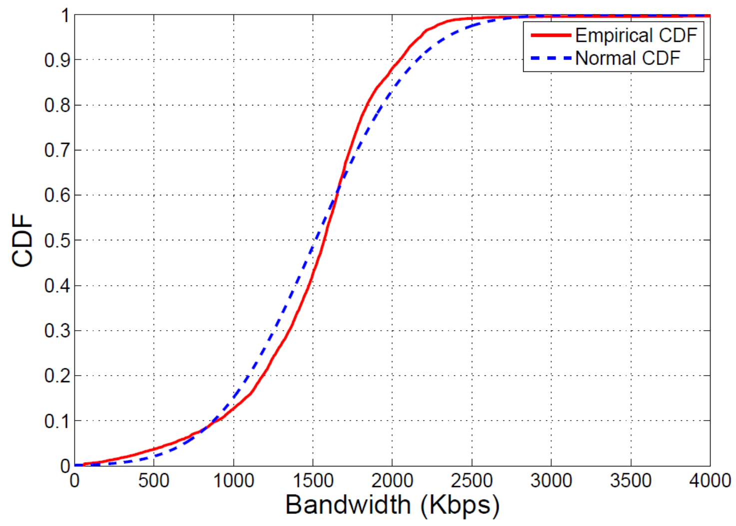
- For a given action (chunk size), state transition probabilities can be calculated if we know the CDF of network bandwidth
  - CDF allows calculation of the probability of a given buffer occupancy (ρ) when the next chunk is downloaded
  - ρ, together with the action (quality level decision), defines the next state S(ρ,q)
- Transition probabilities will change if we have a different CDF
- Different CDFs will lead to different MDP strategies

Approaches to reduce MDP overhead

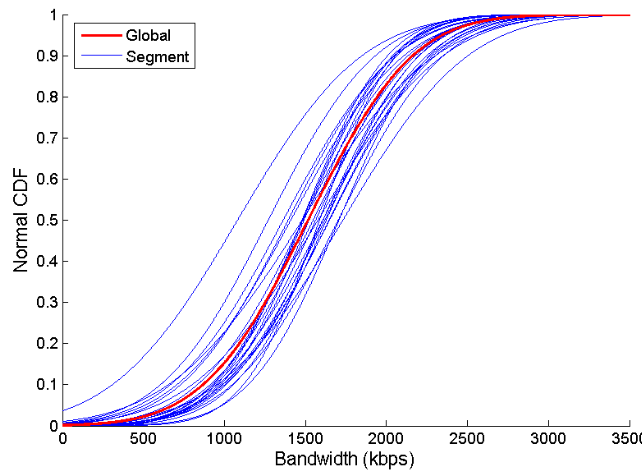
- k-MDP
  - Uses only online data to learn and estimate bandwidth CDF (no offline network measurement)
  - Updates CDF with every sample (each chunk download gives one sample), but recomputes MDP only every k chunk downloads
  - Smaller k means more precise MDP, but higher overhead, and vice versa
- s-MDP
  - Solves MDP with offline network measurement (no online MDP computation)
  - One MDP computed for a large region and used at all locations during a given trip
  - Single MDP strategy to maintain for a large area, but CDF may not be accurate for some locations
- x-MDP
  - One MDP per road segment, but they are all solved offline (no online MDP computation)
  - More precise CDF estimation for every location visited during a given streaming session
  - Requires availability of more extensive bandwidth measurements (measurement needed for every road segments of a city-wide road network)

Simulation Setup

- Big Buck Bunny encoded into 5 different bitrates (qualities) and segmented into 2-sec chunks
- Mobility and 3G bandwidth traces from actual driving in Sydney, Australia
  - 24 Km (22-30 minutes) repeated 70 times
  - Bandwidth measured at 10 sec interval
  - Useful data set to obtain CDF per road segment
  - Normal approx. appears close to empirical CDF (simplifies CDF storage and processing)



	time	latitude	longitude	bandwidth (Kbps)
1	1186549400	-33.919785	151.228913	1663.1440
2	1186549410	-33.919635	151.227787	1964.7330
3	1186549420	-33.91958	151.227322	2038.8659
4	1186549430	-33.91958	151.227322	2011.2631
5	1186549440	-33.91953	151.22692	1838.6578
6	1186549450	-33.91905	151.226322	1208.2767



DASH performance as a function of MDP penalties

low penalties

Quality change penalty

deadline miss

quality level

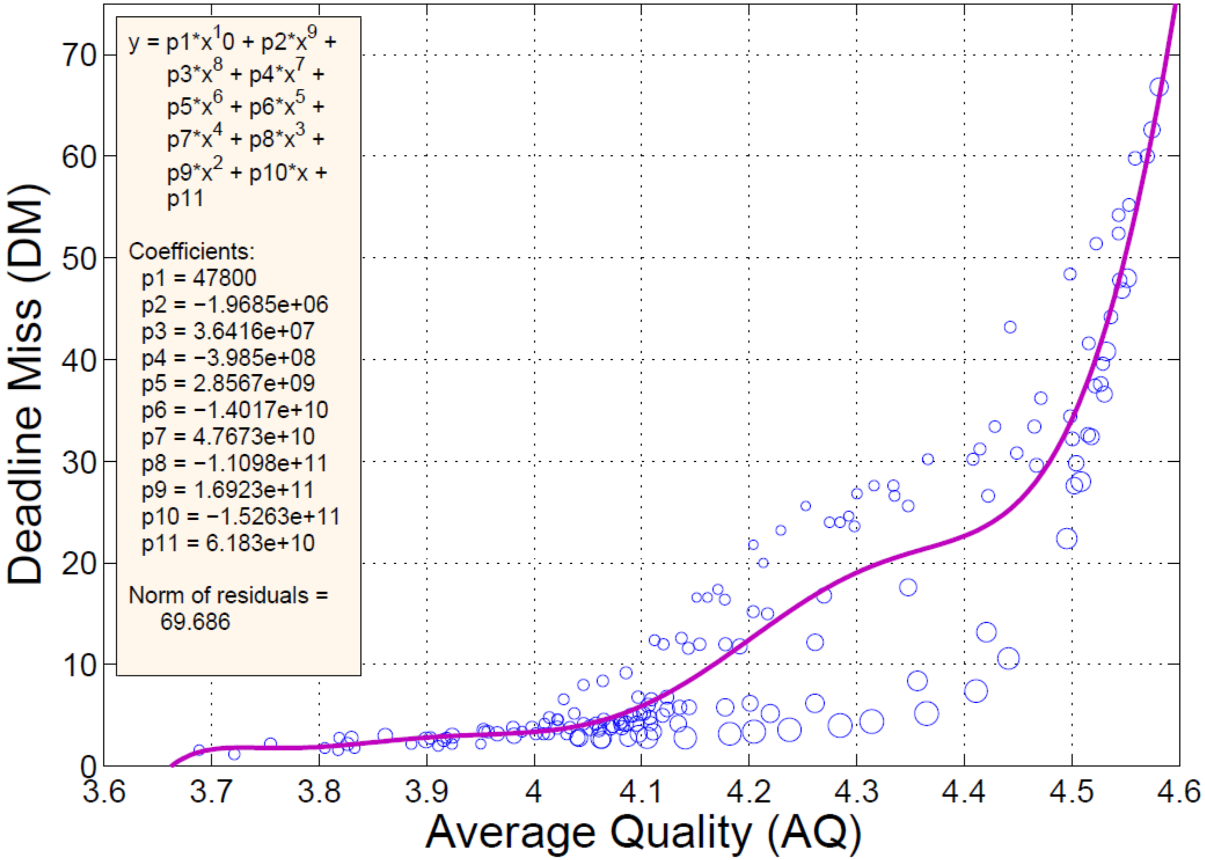
quality change

Deadline penalty

D\C	0.1	0.3	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9
10	195.4 4.74 47.60	202.8 4.74 34.00	207.0 4.75 26.40	230.4 4.77 20.80	268.8 4.80 14.60	277.4 4.81 13.80	280.8 4.81 13.00	282.6 4.82 12.60	282.8 4.81 12.00	286.4 4.81 10.80
15	66.80 4.58 61.40	62.60 4.57 48.80	60.00 4.57 39.80	59.80 4.56 34.20	55.20 4.55 32.00	52.40 4.54 31.20	54.20 4.54 30.40	51.40 4.52 29.20	48.40 4.50 27.40	43.20 4.44 24.60
20	48.00 4.55 64.00	46.80 4.55 46.20	47.80 4.54 39.60	44.20 4.54 34.60	39.60 4.53 32.40	41.60 4.52 32.00	36.20 4.47 28.40	33.40 4.43 26.00	30.20 4.37 23.00	27.60 4.32 20.60
24	40.80 4.53 65.60	36.60 4.53 47.20	37.60 4.53 40.20	37.40 4.52 37.20	34.40 4.50 33.20	30.80 4.45 28.80	31.20 4.41 27.20	26.60 4.34 21.60	26.80 4.30 20.20	25.60 4.25 17.00
27	28.00 4.51 73.40	32.40 4.52 47.20	32.60 4.51 41.80	32.20 4.50 37.20	33.40 4.47 33.60	30.20 4.41 26.80	27.60 4.33 23.60	24.60 4.29 20.00	24.00 4.27 20.20	20.00 4.21 16.20
30	22.40 4.50 78.80	27.60 4.50 51.20	29.80 4.50 44.80	29.60 4.47 36.40	26.60 4.42 30.40	25.60 4.35 26.00	23.60 4.30 21.60	24.00 4.28 20.80	23.20 4.23 17.80	21.80 4.20 16.20
50	10.60 4.44 87.20	13.20 4.42 72.60	17.60 4.35 54.80	16.80 4.27 42.60	15.00 4.22 28.00	15.20 4.20 26.40	16.40 4.18 21.60	17.40 4.17 18.40	16.60 4.16 17.80	16.60 4.15 16.20
70	7.40 4.41 98.20	8.40 4.36 73.80	12.20 4.26 52.40	11.80 4.19 42.80	12.00 4.18 33.60	12.00 4.15 29.20	11.60 4.14 26.20	12.60 4.14 24.60	12.00 4.12 23.20	12.40 4.11 22.40
100	5.20 4.37 107.2	6.20 4.26 65.40	6.20 4.20 47.00	5.80 4.14 39.60	6.80 4.12 35.60	6.60 4.11 33.60	6.80 4.10 28.80	9.20 4.09 26.80	8.40 4.06 26.00	8.00 4.05 23.40
130	4.40 4.31 107.2	5.20 4.22 62.80	5.80 4.14 41.20	5.00 4.12 36.80	5.20 4.10 36.00	5.00 4.09 32.20	4.60 4.08 27.60	5.20 4.04 25.20	4.60 4.02 23.40	6.60 4.03 22.40
150	4.00 4.28 107.0	5.80 4.18 57.20	5.60 4.12 40.60	6.00 4.11 38.20	5.00 4.09 36.20	4.20 4.08 33.00	4.20 4.05 29.20	4.80 4.01 24.40	4.20 4.01 23.00	4.60 4.02 23.80
...	...	...	...	...	...	...	...	...	...	...

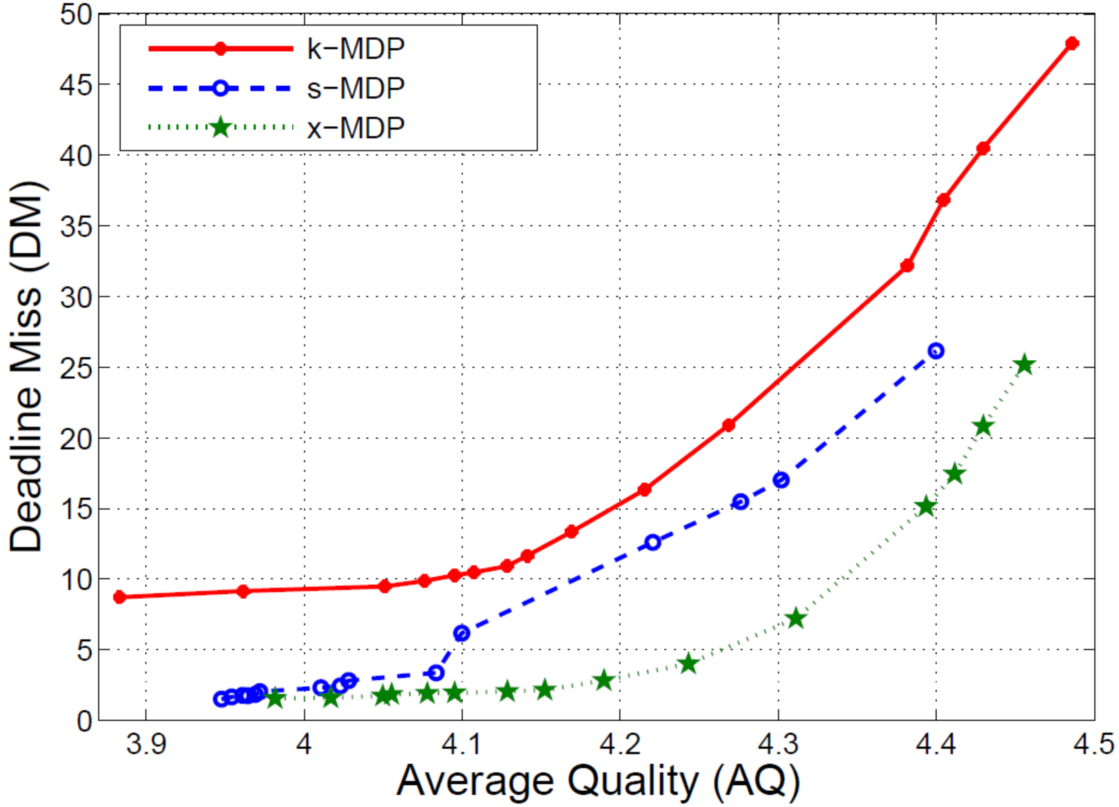
high penalties

MDP tuning allows trade off between quality and deadline miss



Performance of k-MDP with k=1

Comparing k-MDP, s-MDP, and x-MDP

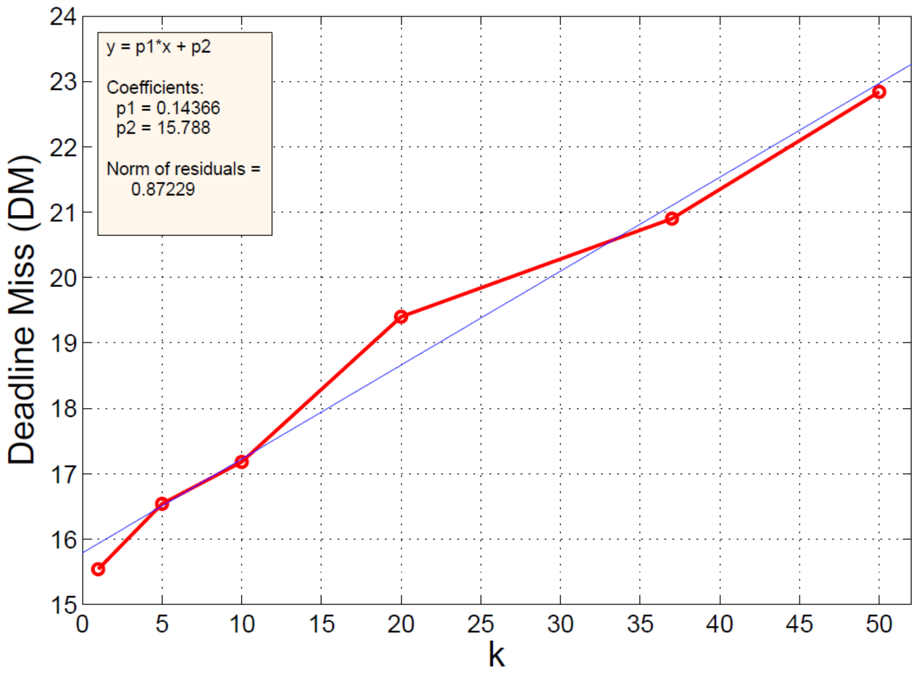


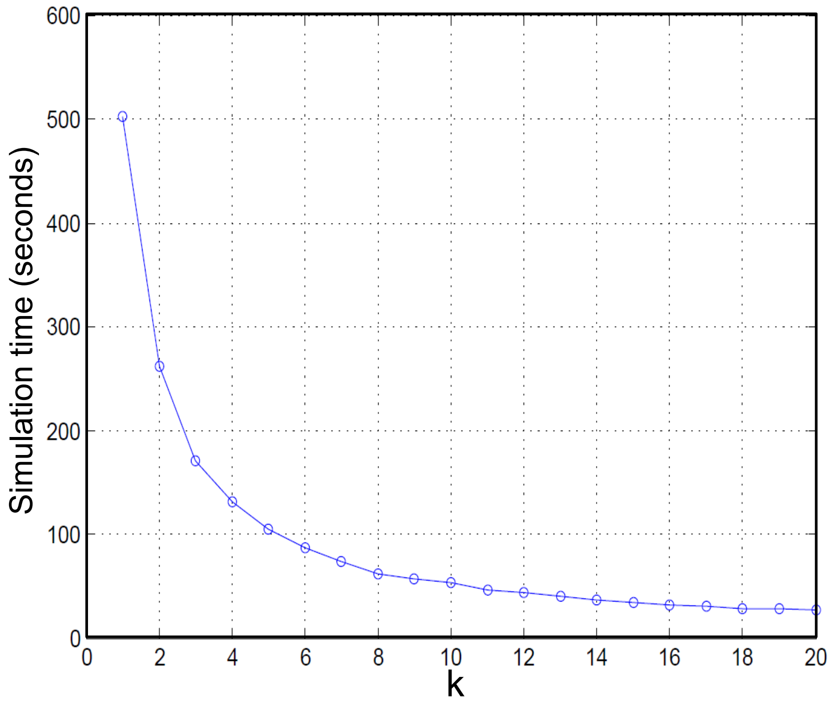
x = 1000 meter

k = 37 (about 37 chunks downloaded per 1000m on average)

k-MDP's performance as a function of k

Deadline miss increases linearly with k





Online computation overhead is very high for k=1, but it can be reduced by a factor of k for k>1

Conclusion

- Although MDP has the potential to improve DASH performance, it increases decision making overhead in a mobile device
- We have explored 3 different implementations of MDP with a goal of reducing decision making overhead during a streaming session
- k-MDP reduces online optimization overhead of MDP by a factor of k
  - but increases deadline miss linearly as a function of k
  - it allows DASH decision making with no offline network measurements, but performs worse than offline optimization
- s-MDP and x-MDP can be used to solve MDP offline, eliminating the online computation overhead completely
  - but they require offline network measurement
- x-MDP outperforms s-MDP due to location-sensitivity of bandwidth distributions
  - but it requires more extensive offline network measurement and maintenance of a large number of MDP strategies for a given road network (one per road segment)