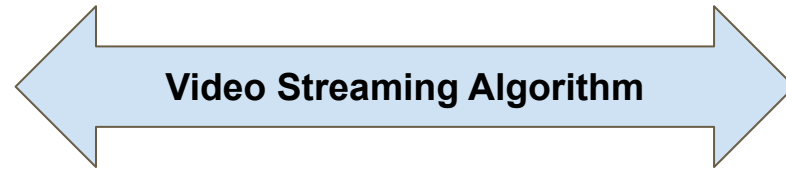

Pitfalls of data-driven networking: A case study of latent causal confounders in video streaming

— P. C. Sruthi, Sanjay Rao, Bruno Ribeiro —



Say you want design a video streaming system...

Say you want design a video streaming system...



What If...

- A different algorithm had been used?
- Viewers started playing 4K videos? Would they experience buffering?

What If...

- A different algorithm had been used?
- Viewers started playing 4K videos? Would they experience buffering?

Counterfactual questions

What this talk is about

- What are the challenges involved in answering counterfactual questions for networked systems?

What this talk is about

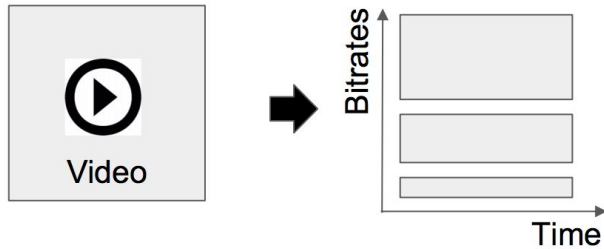
- What are the challenges involved in answering counterfactual questions for networked systems?
- A study of these challenges in the context of video streaming algorithms

What this talk is about

- What are the challenges involved in answering counterfactual questions for networked systems?
- A study of these challenges in the context of video streaming algorithms
- Limitations of current methods, and a preliminary approach to overcome these challenges

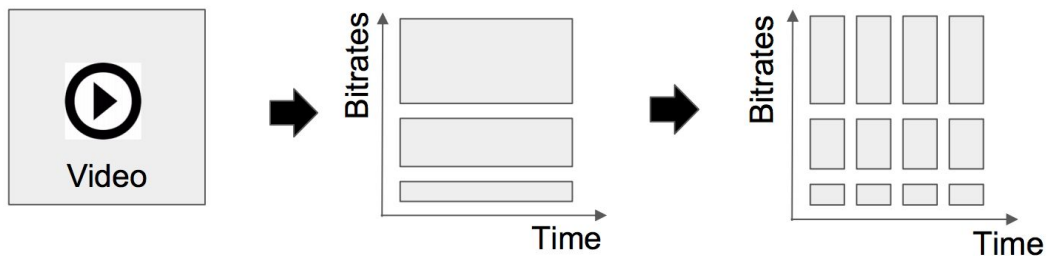
Background: Video Streaming (ABR)

Background: Video Streaming (ABR)



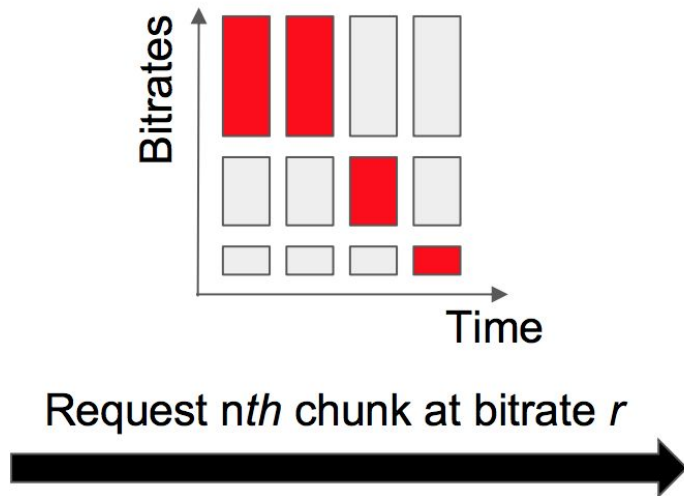
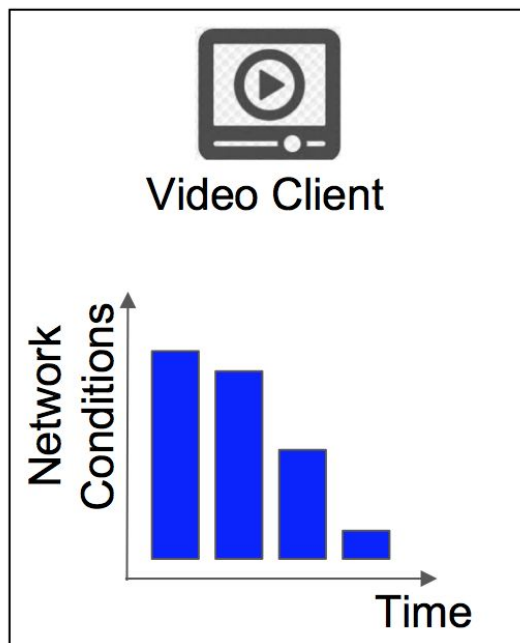
A video is encoded into multiple qualities (bitrates)

Background: Video Streaming (ABR)



Each bitrate is split into chunks

Background: Video Streaming (ABR)



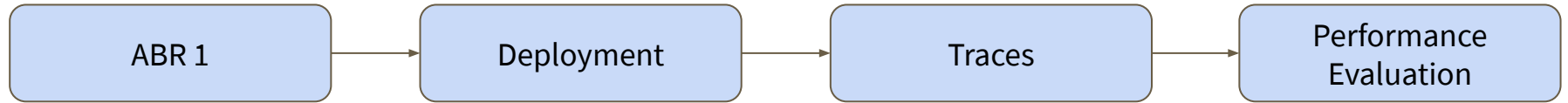
Counterfactuals for video streaming

- What if ABR 2 had been used instead of ABR 1?

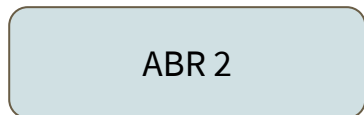
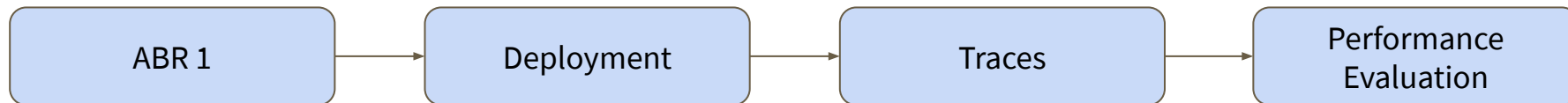
Counterfactuals for video streaming

- What if ABR 2 had been used instead of ABR 1?
- *Alternatively*, what if a different sequence of bitrates had been downloaded?

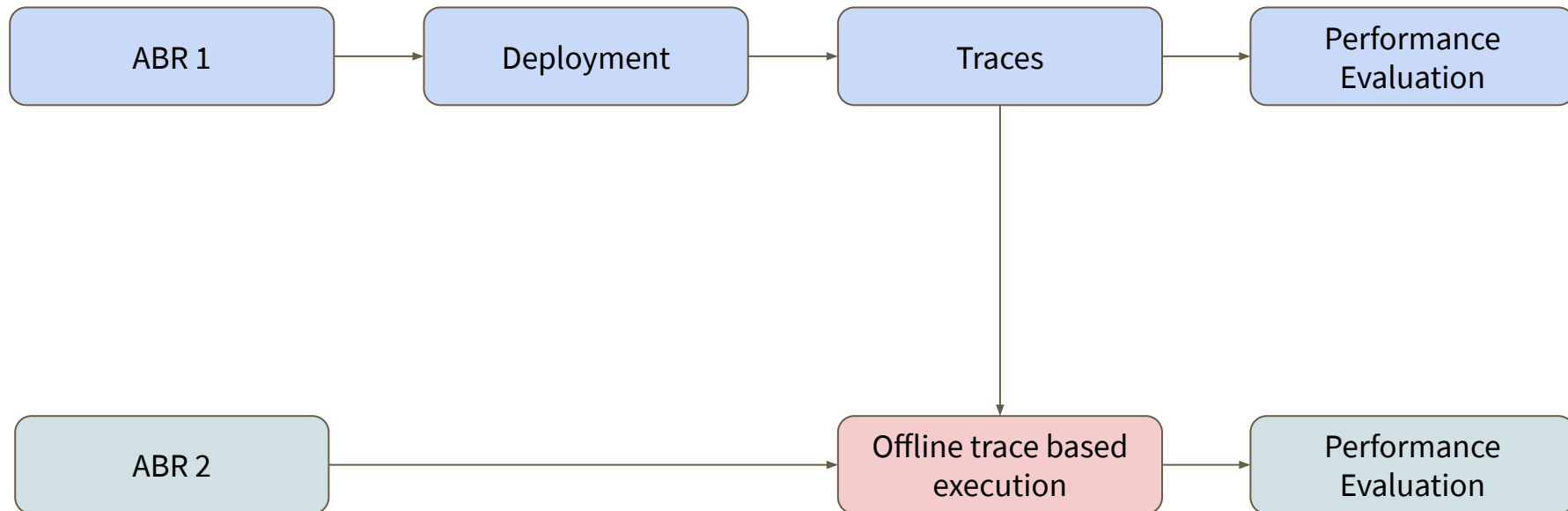
Counterfactuals for video streaming



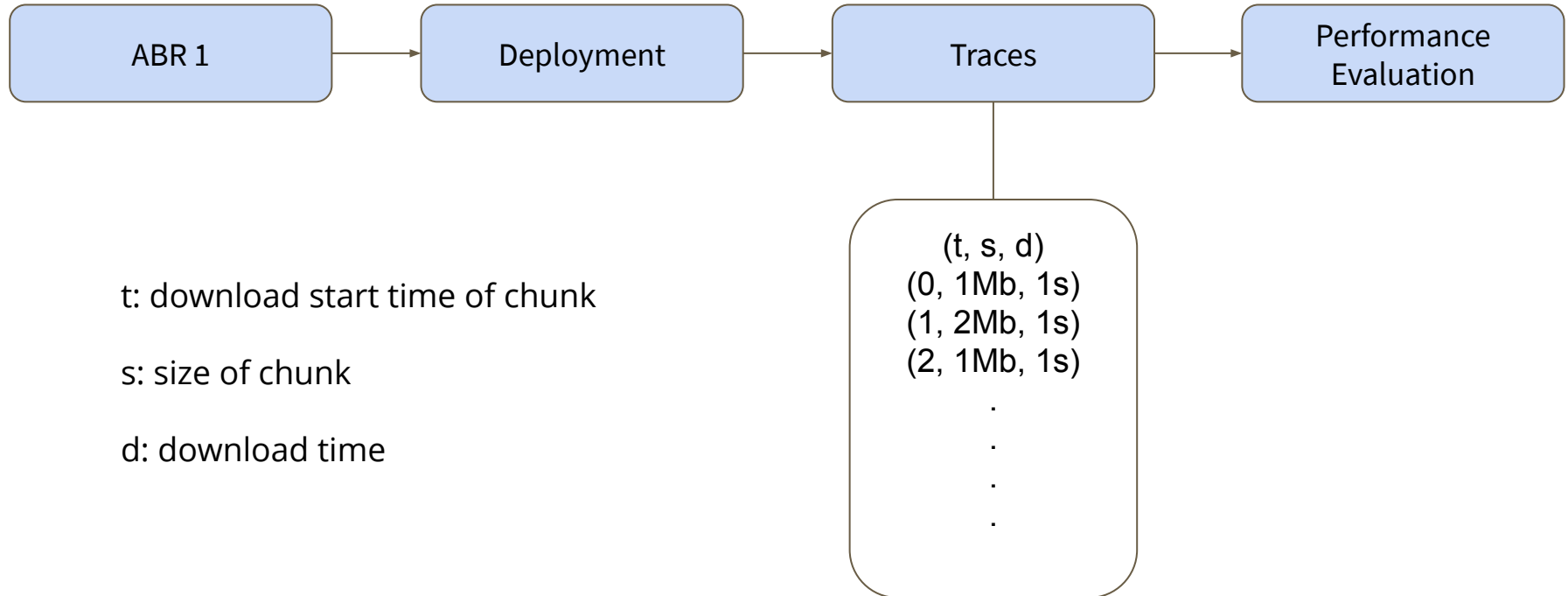
Counterfactuals for video streaming



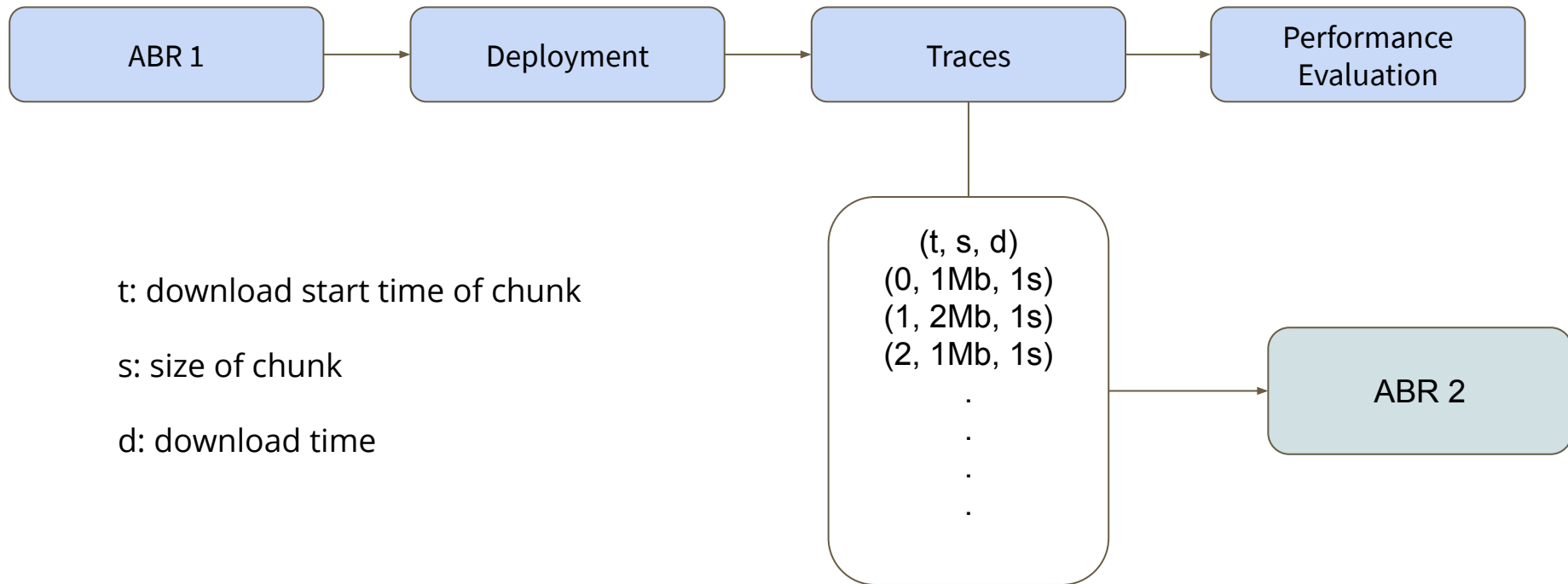
Counterfactuals for video streaming



Evaluating video streaming systems using traces



Evaluating video streaming systems using traces

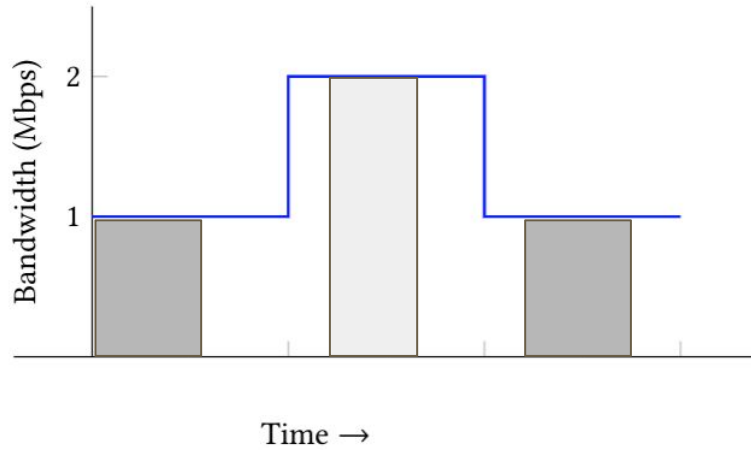


What can go wrong with using traces?

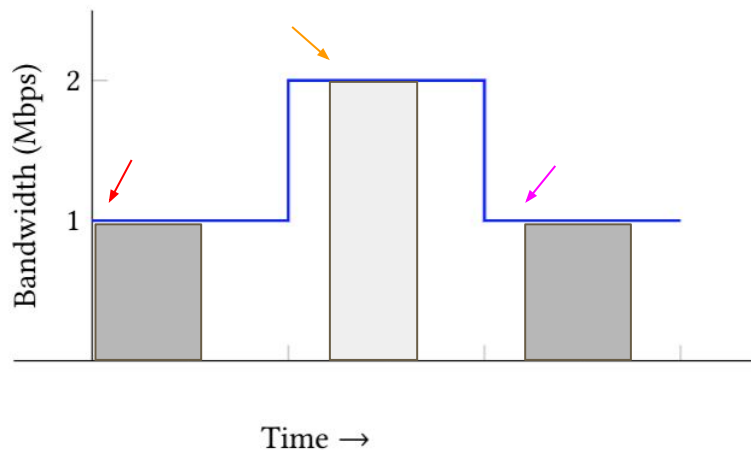
What can go wrong with using traces?

- *Traces generated by adaptive algorithms can affect trace driven evaluation!*

What can go wrong with using traces?

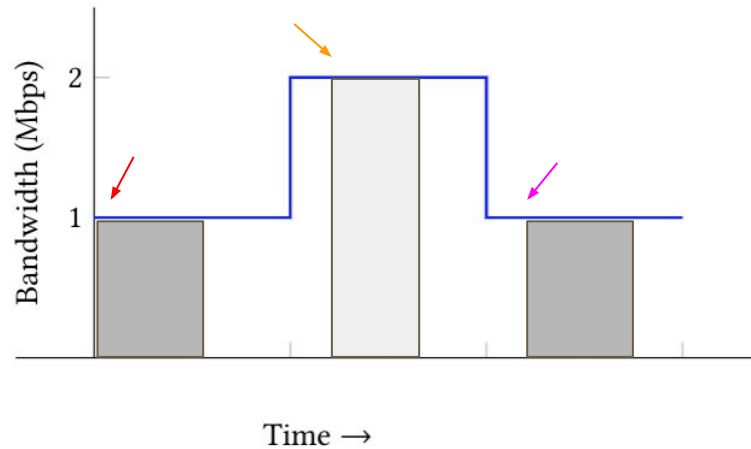


What can go wrong with using traces?

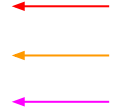


- **ABR-Probe** probes bandwidth before downloading a chunk
- Chooses bitrate to match the probed bandwidth

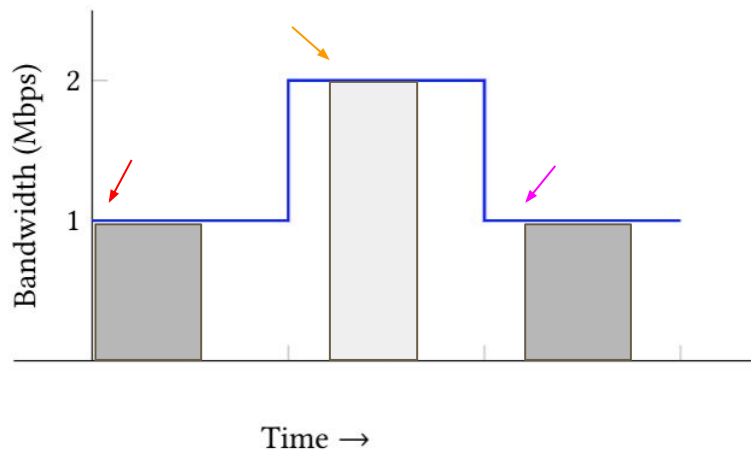
What can go wrong with using traces?



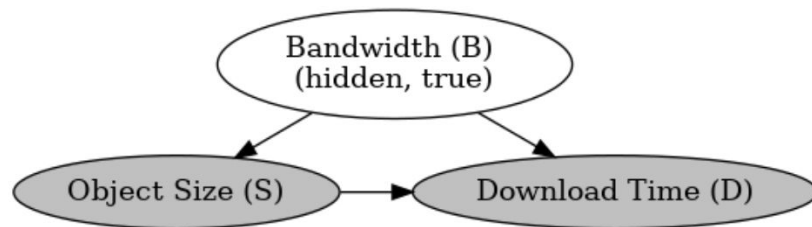
(t, s, d)
(0, 1Mb, 1s)
(3.2, 2Mb, 1s)
(4.6, 1Mb, 1s)
.
.
.
.



What can go wrong with using traces?

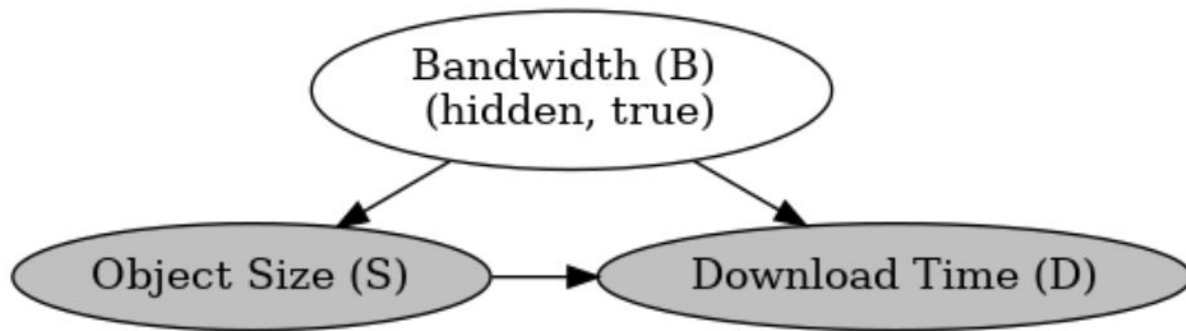


(t, s, d)
(0, 1Mb, 1s)
(3.2, 2Mb, 1s)
(4.6, 1Mb, 1s)
.
.
.
.



The issue of confounders

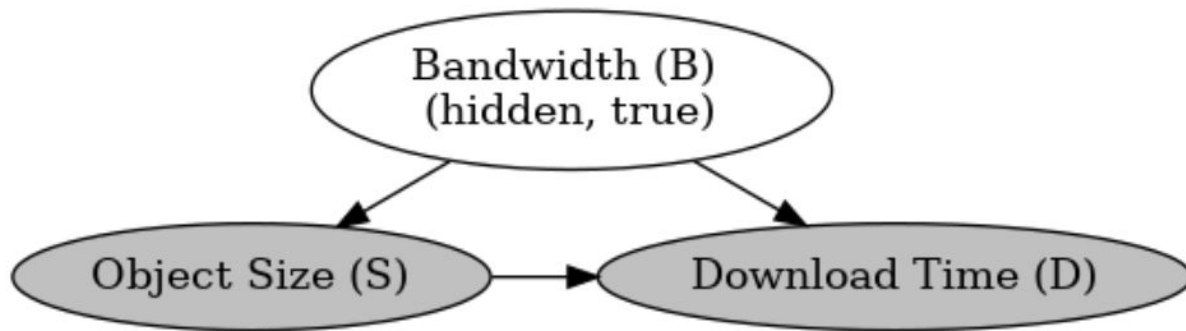
Confounders induce dependencies in the data that are often unaccounted for.



The issue of confounders

Confounders induce dependencies in the data that are often unaccounted for.

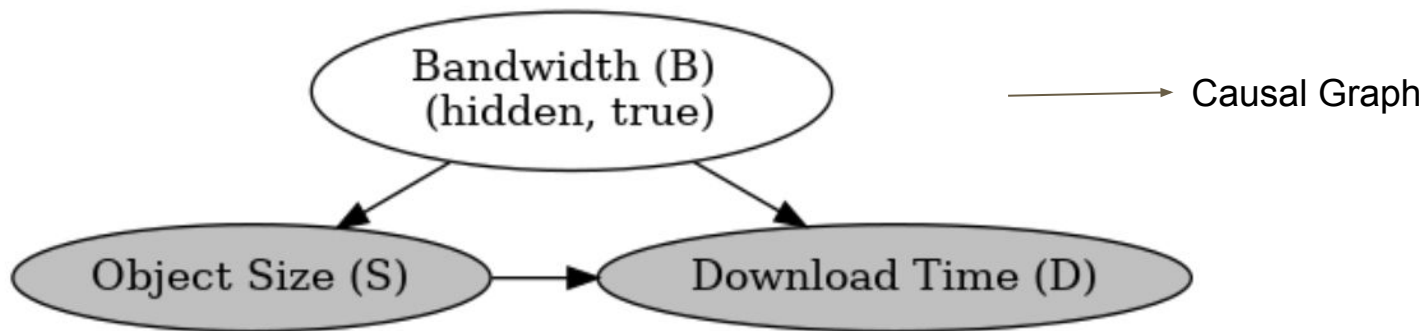
This can affect the accuracy of trace based execution.



The issue of confounders

Confounders induce dependencies in the data that are often unaccounted for.

This can affect the accuracy of trace based execution.



Existing approaches to deal with confounders

- Randomized Controlled Trials (RCTs)
 - Choose the bitrates at random so that the bandwidth doesn't affect it
 - RCTs don't work here - Trace collection is impractical, other data dependencies

Existing approaches to deal with confounders

- Randomized Controlled Trials (RCTs)
 - Choose the bitrates at random so that the bandwidth doesn't affect it
 - RCTs don't work here - Trace collection is impractical, other data dependencies

- Observational Studies (Matching on confounders)
 - Find data in the original trace that matches what you'd like to estimate in your new system, and use that as a measurement
 - Do not account for latent confounders [1][2]

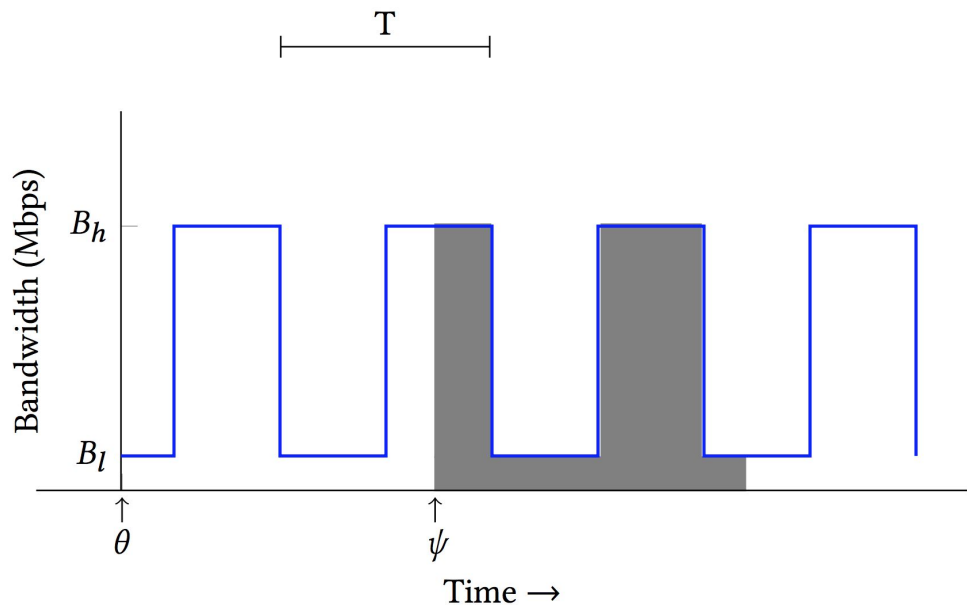
[1] S. Shunmuga Krishnan and Ramesh K. Sitaraman. 2012. Video stream quality impacts viewer behavior: inferring causality using quasi-experimental designs. In Proceedings of the 2012 Internet Measurement Conference (IMC '12)

[2] Detecting network neutrality violations with causal inference. In Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies

What if you could account for latent confounders?

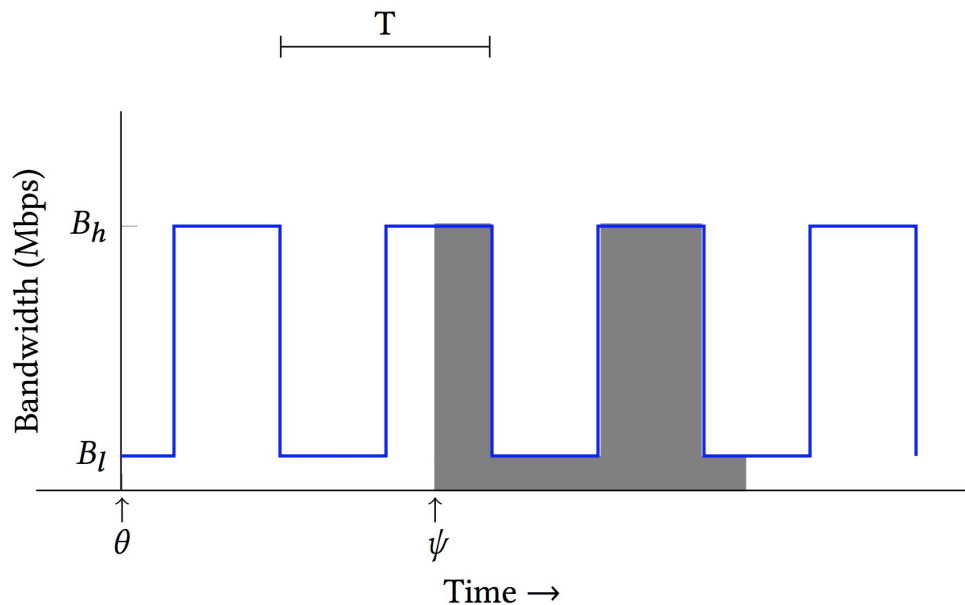
- We conducted a case study on the simplest scenario that illustrated the problem

Illustrative Case Study



- Create trace by downloading a video using ABR-Probe
- Use trace to evaluate performance of second bitrate sequence

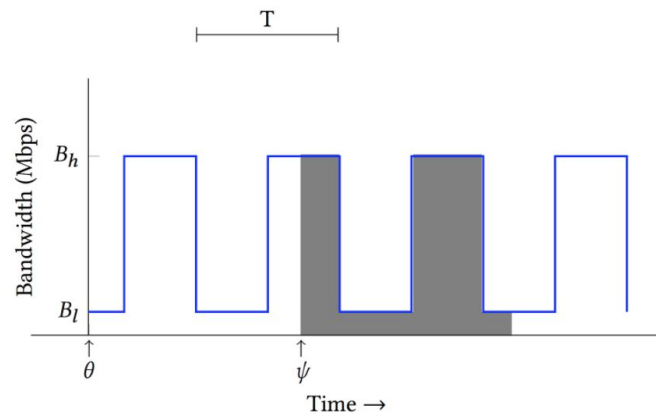
Illustrative Case Study



- Create trace by downloading a video using ABR-Probe
- Use trace to evaluate performance of second bitrate sequence
- Assumptions:
 - θ : session phase, hidden
 - ψ : chunk start phase, hidden
 - B_h, B_l, T are known

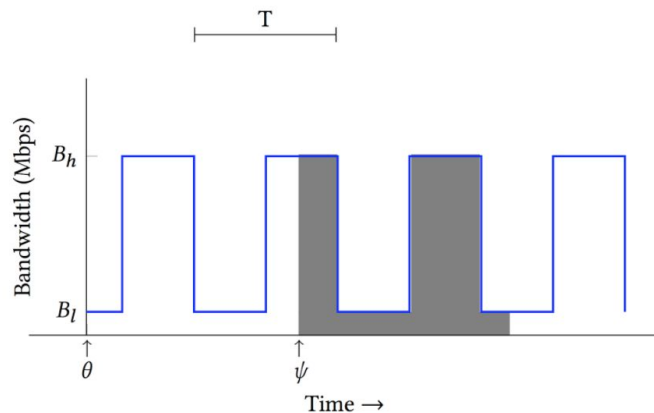
Our Approach

- Construct causal graph for trace production process
- Infer hidden confounders from the data



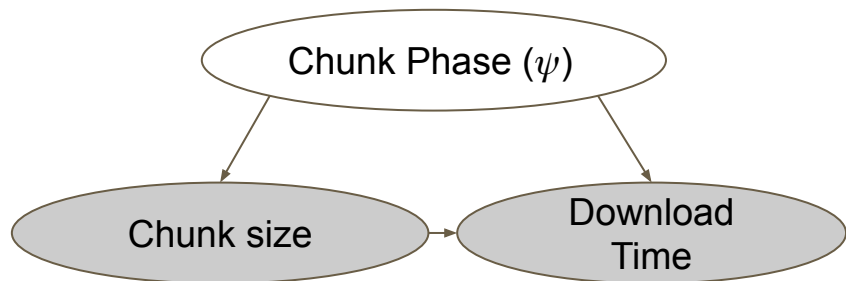
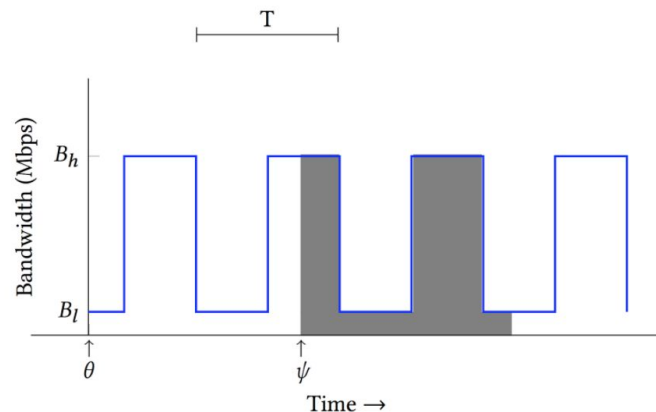
Our Approach

- Construct causal graph for trace production process
- Infer hidden confounders from the data
- Use trace with inferred confounders to evaluate performance of second sequence



Our Approach

- Key Idea:
 - *Infer the chunk phase explicitly from the data*
- Use Maximum A Posteriori estimation
 - All of the details in the paper



Evaluation

- **Trace Production:** ABR, Randomized bitrates

Evaluation

- **Trace Production:** ABR, Randomized bitrates
- **Trace based evaluation**
 - Calculate download times of new sequence of bitrates using only the trace as input, with different methods

Evaluation

- **Trace Production:** ABR, Randomized bitrates

- **Trace based evaluation**
 - Calculate download times of new sequence of bitrates using only the trace as input, with different methods

 - **Evaluation metric:** Error in download time calculation from trace vs ground truth deployment

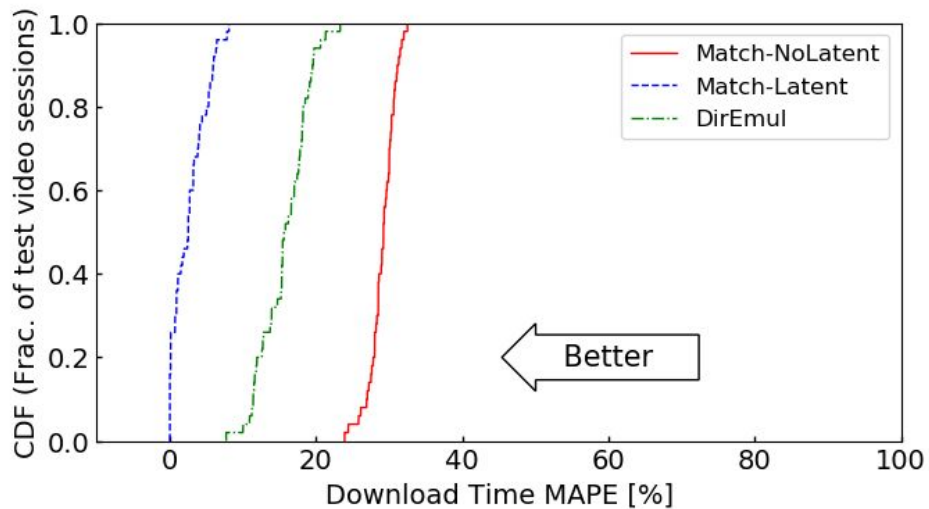
 - How accurate was it in answering the counterfactual compared with ground truth?

Evaluation

- **Trace based evaluation methods**
 - **Direct Emulation** - Use observed throughput from trace as bandwidth model
 - **Match - No Latent** - Match on measured features only (bitrate)
 - **Match - Latent** - Our method: match on bitrate and inferred chunk phase

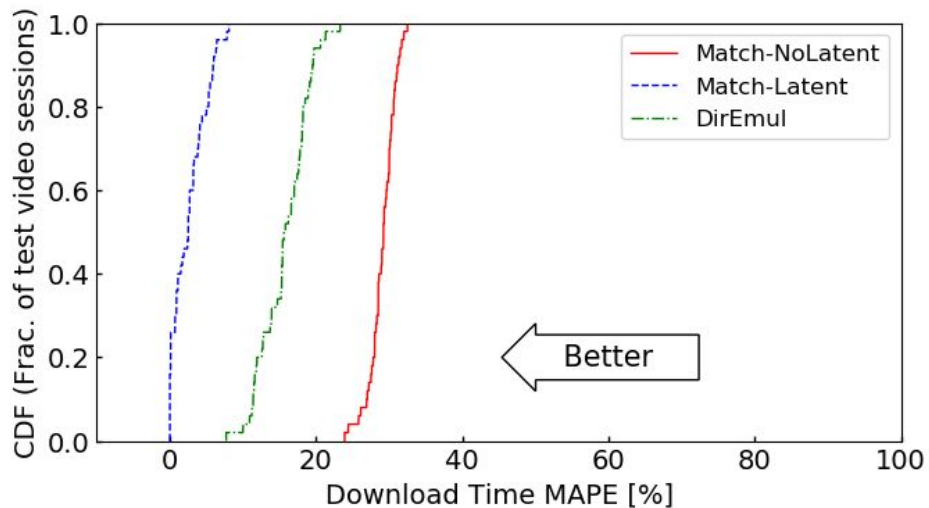
Takeaways

Trace Production: ABR-Probe



Takeaways

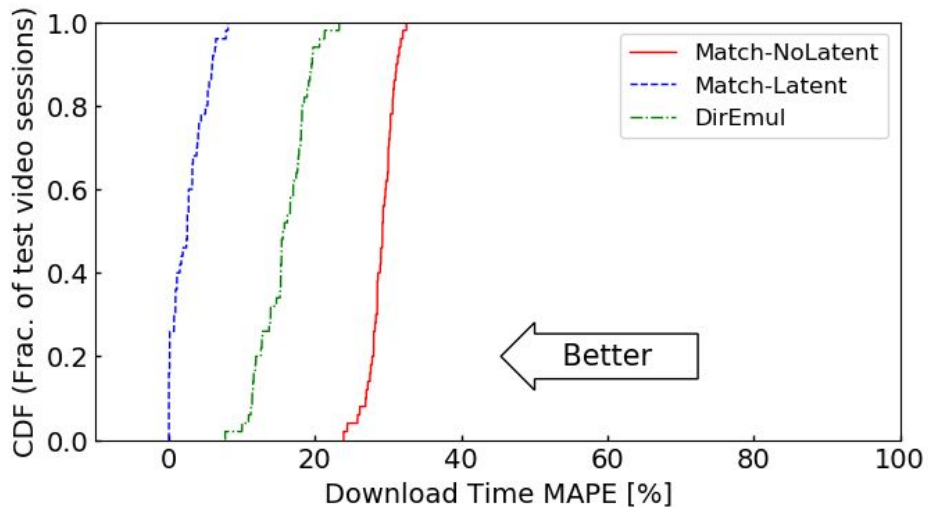
Trace Production: ABR-Probe



- Direct Emulation based on the observed throughputs is not accurate for evaluation - median error ~18%

Takeaways

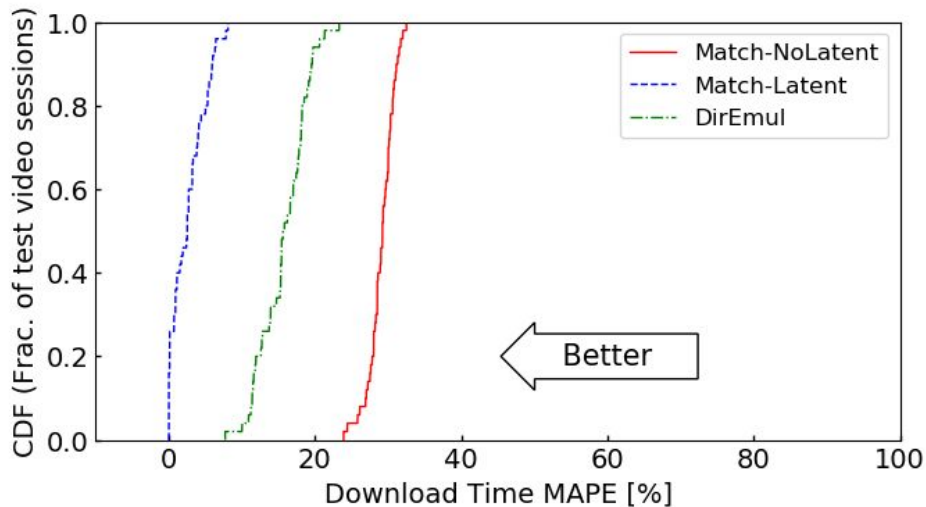
Trace Production: ABR-Probe



- Direct Emulation based on the observed throughputs is not accurate for evaluation - median error ~18%
- Performing matching without accounting for confounders can be even worse

Takeaways

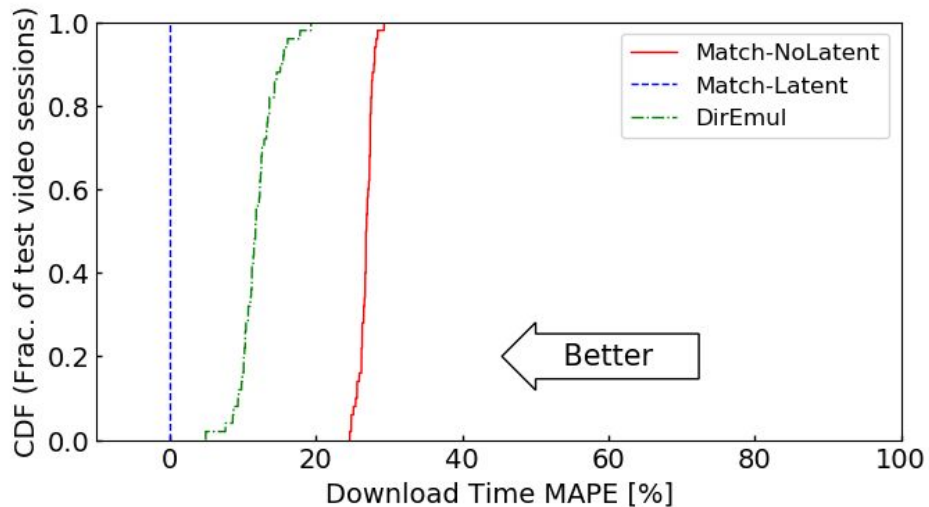
Trace Production: ABR-Probe



- Direct Emulation based on the observed throughputs is not accurate for evaluation - median error ~18%
- Performing matching without accounting for confounders can be even worse
- Matching on latent confounders is the most accurate

Using RCTs

Trace Production: Randomized bitrates



- Similar results
- Match-Latent is optimal

Conclusions and Future Directions

- First step towards answering counterfactual questions with video streaming systems
 - Key challenge: True bandwidth process is not available - latent confounders

Conclusions and Future Directions

- First step towards answering counterfactual questions with video streaming systems
 - Key challenge: True bandwidth process is not available - latent confounders
- Preliminary approach to deal with latent confounders
 - RCTs and matching techniques insufficient without considering latent confounders

Conclusions and Future Directions

- First step towards answering counterfactual questions with video streaming systems
 - Key challenge: True bandwidth process is not available - latent confounders
- Preliminary approach to deal with latent confounders
 - RCTs and matching techniques insufficient without considering latent confounders
- **Challenges and Future Directions:**
 - Generalization towards richer bandwidth processes, what this means for more complex scenarios