Modeling Advertiser Bidding Behaviors in Google Sponsored Search
A Mirror Attention Mechanism

Liang Liu (Google)

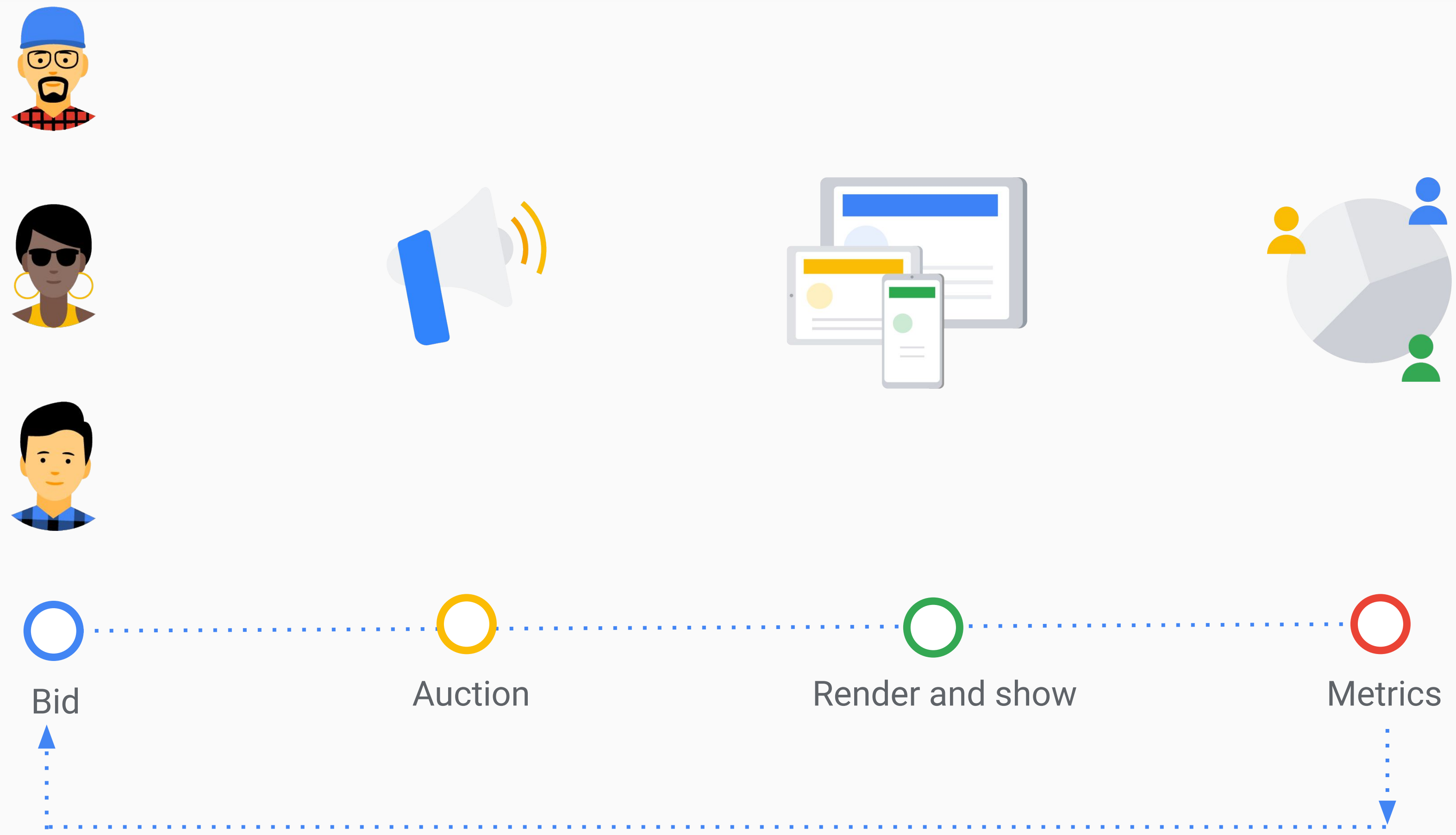
Work by:
Suqi Liu, Liang Liu, Sugato Basu, Jean-François Crespo

Background

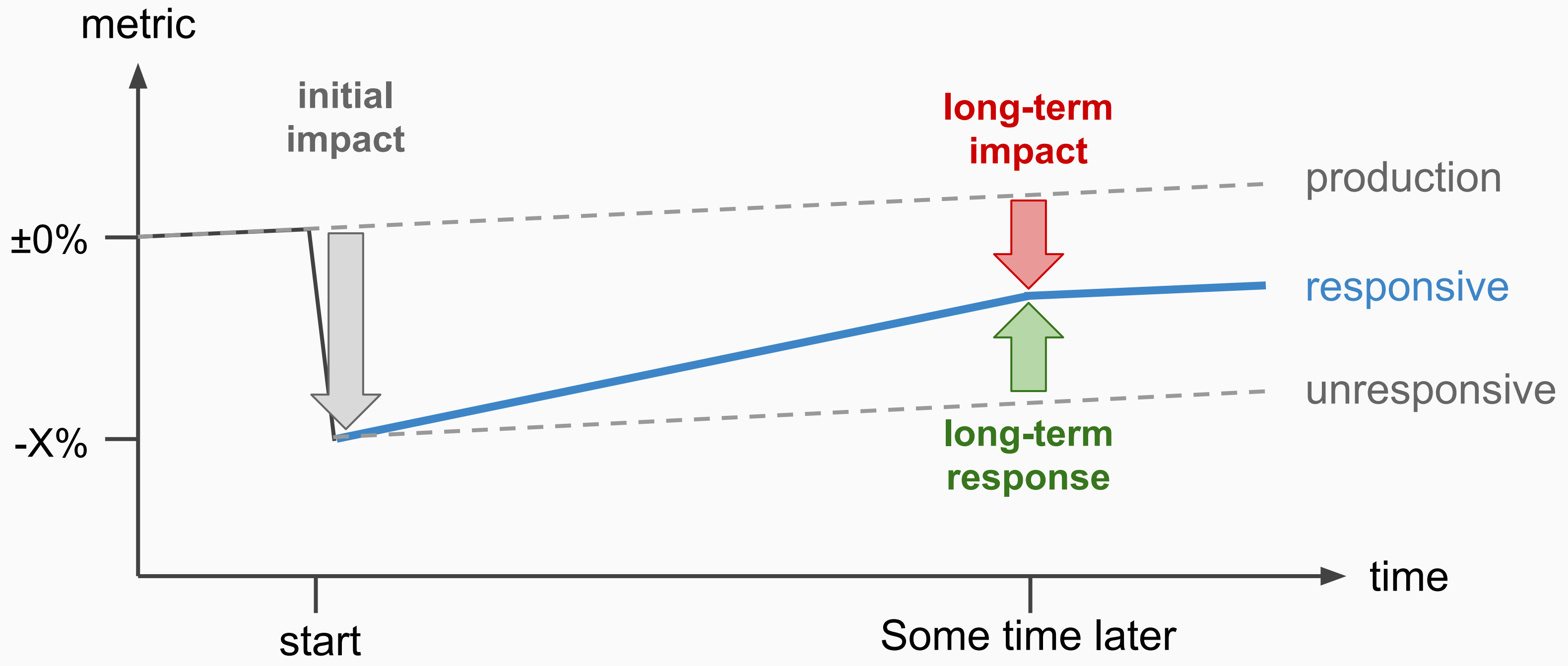
Advertisers respond to feature launches in ads system

- A feature launch can result in changes in certain metrics
- Advertisers respond in various ways to the metrics that they observed
- Long term effect of a launch needs to take these response into account

How do Launches Affect System Metrics

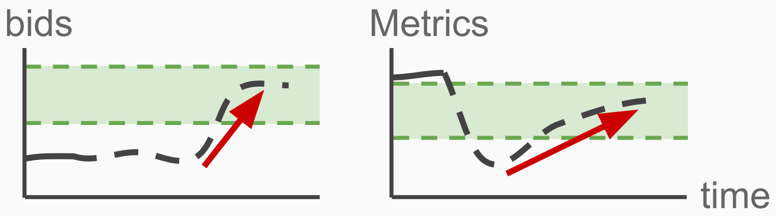


How do Advertisers Respond



Goal of Modeling Advertisers

- Predict metrics considering long-term advertiser response to launches in ads system
- Aims to estimate advertiser response **before / during / after** a launch



How do we model advertiser response?

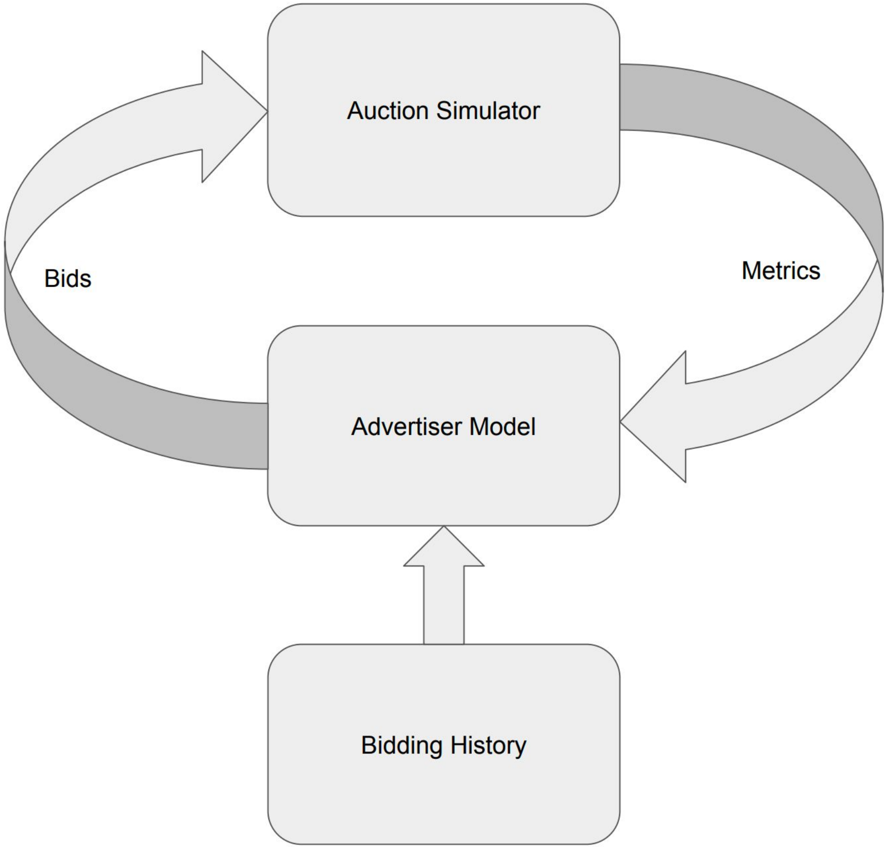
- Advertisers can respond to a change in the ad system in different ways
  - Adjust bid, budget, campaign structure, etc.
- We currently model *bid adjustments* made by advertisers
- Have to model (a) individual response, (b) response interaction via auction

Complexity of Problem

- Advertisers reactions are affected by various reasons
  - E.g., targeting strategy changes
- Advertiser responses are not IID
  - Interaction via the auction in each impression
- Advertiser's reaction can be long-term
  - Change budget allocation at end of quarter
- Super-tricky to get advertiser response ground truth
  - Data sparsity, noise

Advertiser Response Offline Experiment

- Reinforcement-learning like framework
    - Decouple the system(auction) and advertisers.
    - Iteratively run two components
- Treat advertisers as black-box
    - Directly model advertiser response from historical data.
    - Only model short-term response.



Advertiser Response Models

- Descriptive:
  - Invariant models
    - Preserve invariants: Spend / Conversions / Impression/ CPC
  - Other strategies (e.g., constrained utility maximization)

- Predictive:
  - Prediction model for direct regression

Metrics Features and Transformation

Raw features:

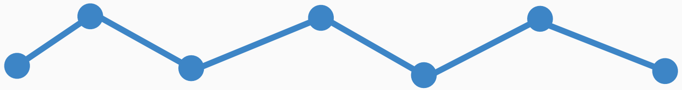
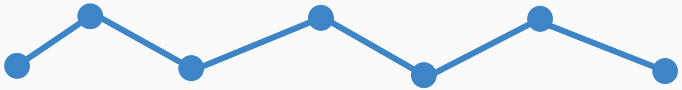
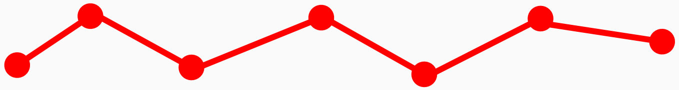
- Impressions
- Clicks
- Conversions
- Budget
- Cost
- Slot

Derived features

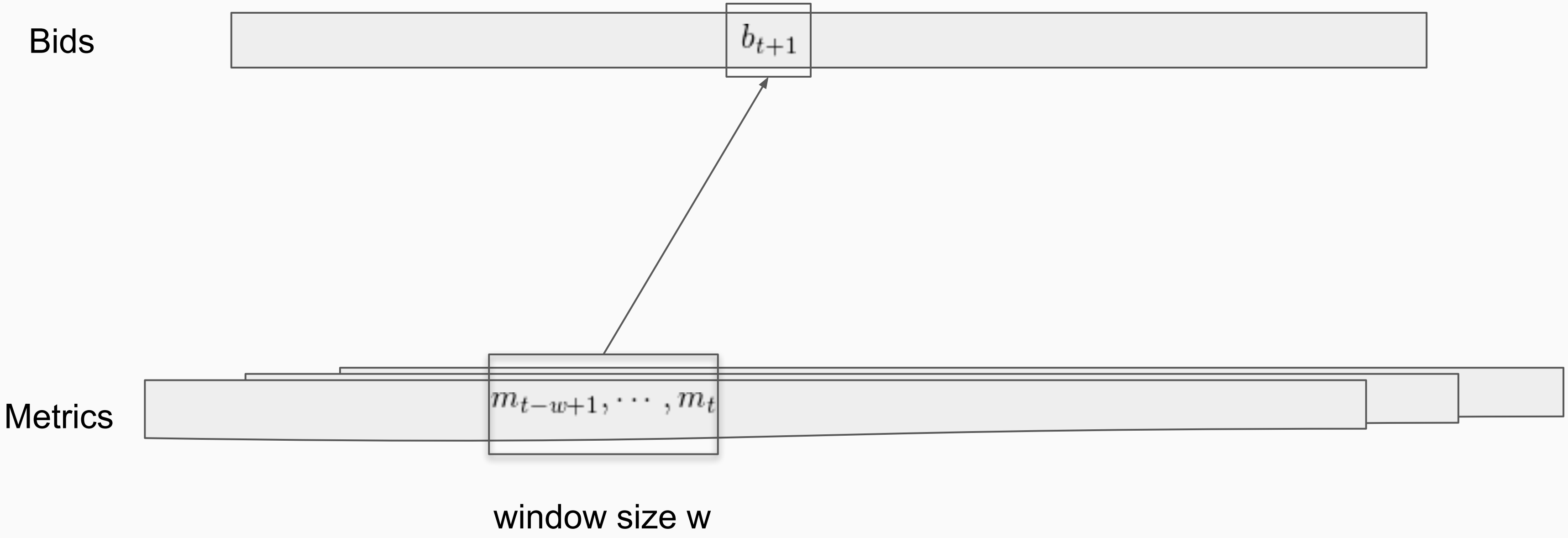
- CTR (clicks/impressions)
- CVR (conversions/clicks)
- CBR (cost/budget)
- CPC (cost/clicks)

https://support.google.com/google-ads/, *snapshot on 2019.07.20*

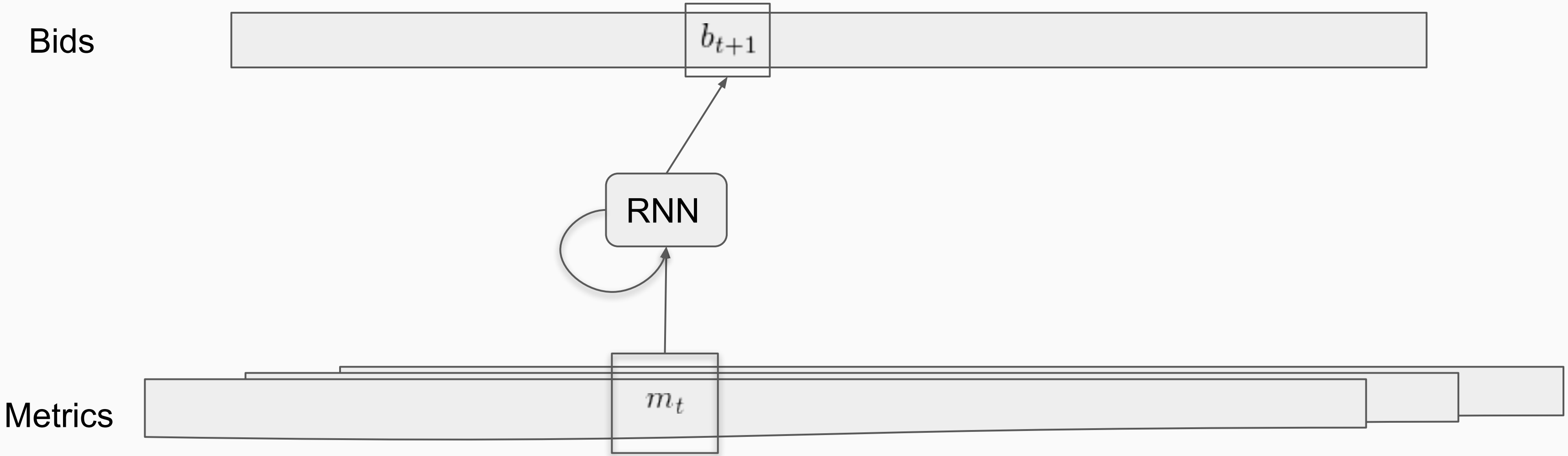
Data Form: Multivariate Time Series

Driver Sequence	Impressions	
	Clicks	
	Cost	
	Ads Positions	
	…….	
Response Sequence	Bids	

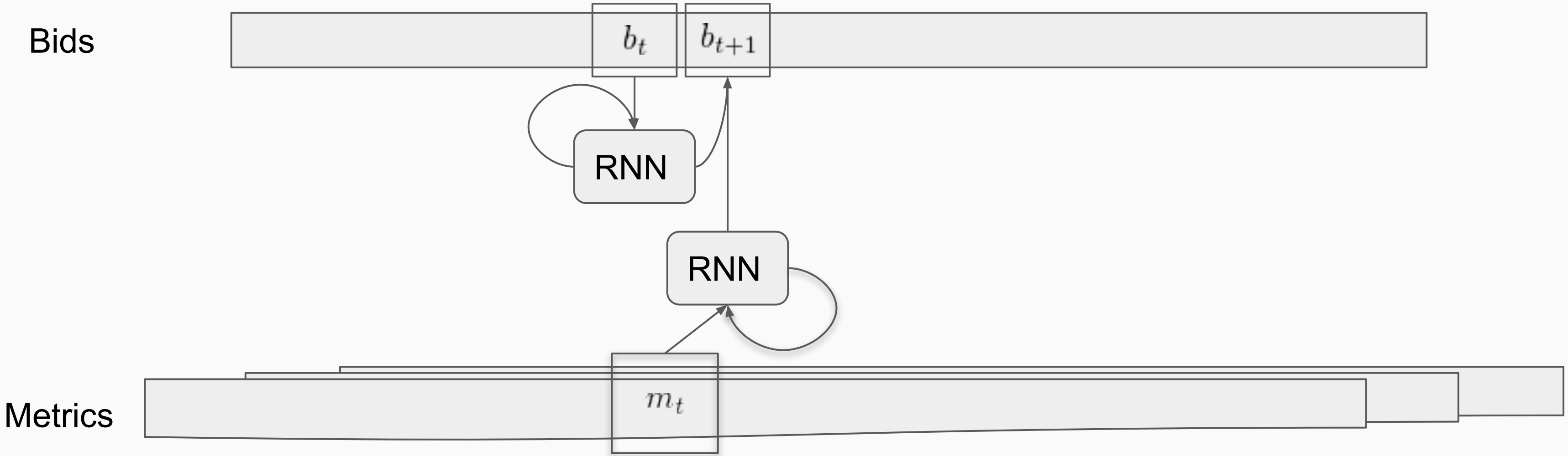
Model Trials: Regression Model



Model Trials: Single Sequence Model(RNN)



Model Trials: Double Sequence Model (Dual RNN)



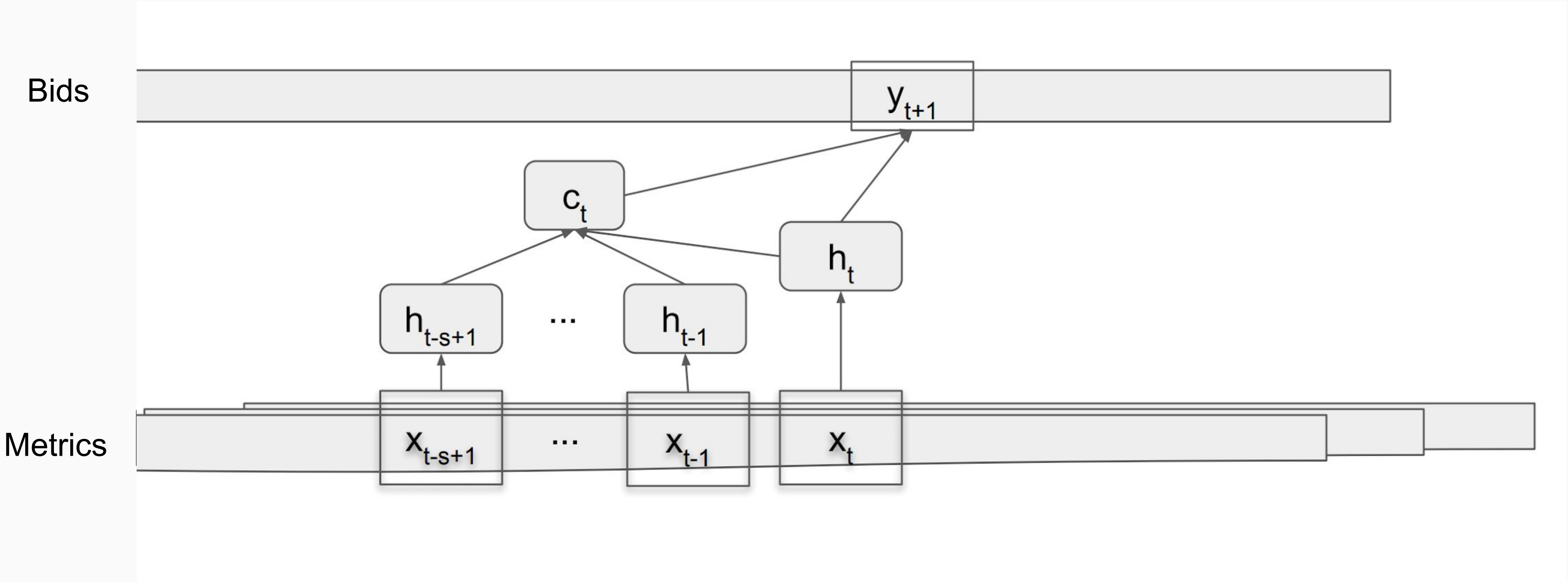
Attention Mechanism

- Ideally sequence models should be able to capture long range dependencies, but is difficult in reality.
- When making prediction, focus (i.e., attend) on relevant part of input

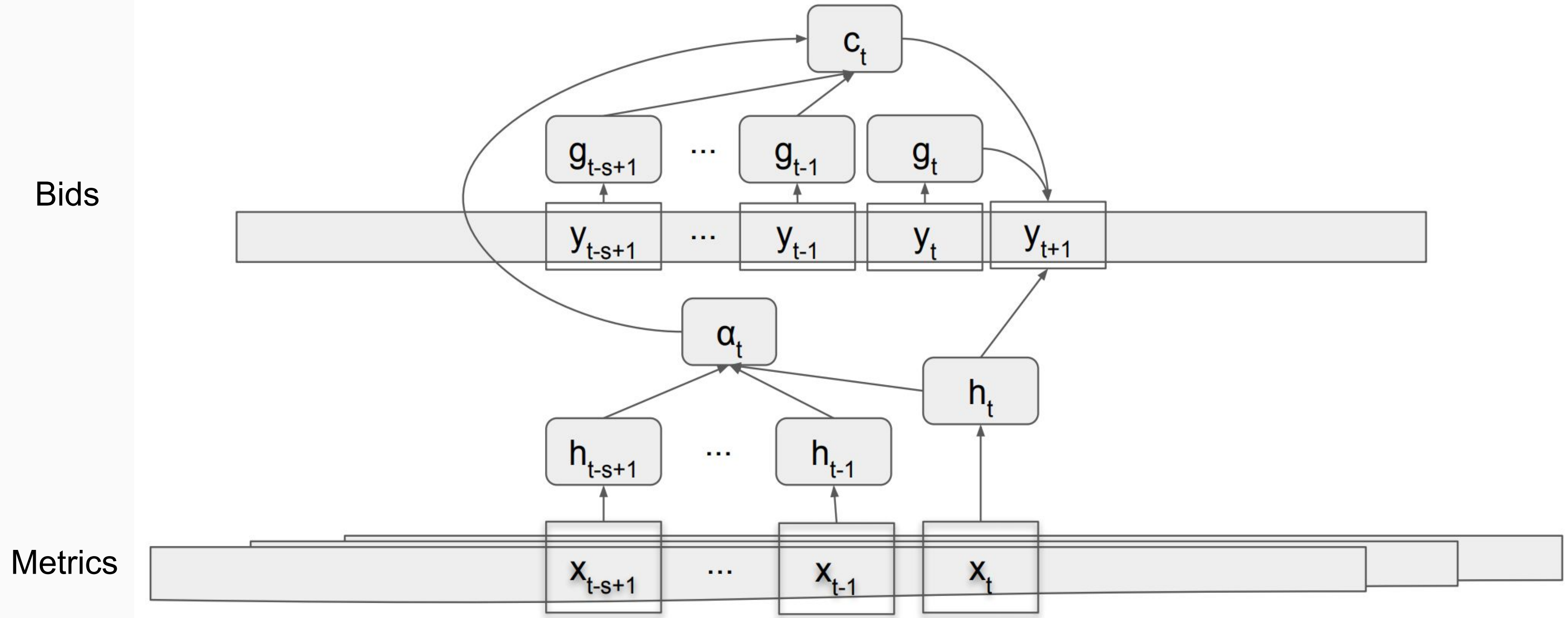
A woman is throwing a frisbee in a park.

- In our context, to focus on ***relevant parts of historical sequence***

Model Trials: Casual Attention Model



Model Trials: Mirror Attention Model



Testbed: Air Quality Data

- Air quality data from UCI ML repository [source]
- Multivariate time-series
- The dataset resembles the advertiser response
    - The concentration of pollutant has its own evolution [response metrics]
    - Concentration is influenced by weather conditions like temperature, pressure, wind speed, cumulative hours of rain, etc [driver metrics]

Results on Air Quality Test Data

	3 hours	6 hours	12 hours	24 hours
NAIV	100.00	100.00	100.00	100.00
REGR	111.11	121.98	113.65	75.63
SSEQ	**87.70**	89.03	73.28	66.41
CATT	94.36	**81.44**	73.32	67.99
DSEQ	93.82	84.47	**72.72**	**65.21**
MATT	**89.89**	**81.82**	**72.82**	**65.67**

- DSEQ and MATT achieves better results when we increase the difficulty of the prediction task with larger predicting gap
- MATT performs consistently among the best models

Results on Advertiser Bid History Data

Model	Relative mean squared error
NAIV	100.00
REGR	85.62
SSEQ	80.82
CATT	74.11
DSEQ	64.96
MATT	**61.98**

- Length of attention window plays an important role
- The dimensions of hidden states in the driver sequence and response sequence significantly contribute to performance
- Parameter tuning discussed in paper

Conclusion

- Introduced a new data-driven approach to advertiser bid prediction
- A novel mirror attention mechanism tailored to the sequential prediction task was proposed
- The first step in our attempts towards understanding advertiser behaviors via sequence modeling
- Following up work to introduce more auction rules and policy into the models to strengthen from a pure multivariate time series model

Beyond Bid Response: Other Applications

The model we developed can potentially have more impacts when applied to the following tasks.

- Resource usage in systems
- User behavior modeling
- Weather prediction
- Financial market forecasting

Thanks!