### Online Interval Scheduling with Predictions



Joan Boyar<sup>1</sup>, Lene M. Favrholdt<sup>1</sup>, Shahin Kamali<sup>2</sup> and Kim S. Larsen<sup>1</sup>

July 31, 2023

1. University of Southern Denmark, Odense, Denmark 2. York University, Toronto, Canada

The Algorithms and Data Structures Symposium (WADS)



#### • Interval Scheduling and Disjoint Path Allocation problems.



- Interval Scheduling and Disjoint Path Allocation problems.
- Online Algorithms with Predictions



#### • Interval Scheduling and Disjoint Path Allocation problems.

- Online Algorithms with Predictions
- Main results:
  - Disjoint Path Allocation problem
  - Interval Scheduling: competitive results, consistency/robustness tradeoffs, and experimental results



#### • Interval Scheduling and Disjoint Path Allocation problems.

- Online Algorithms with Predictions
- Main results:
  - Disjoint Path Allocation problem
  - Interval Scheduling: competitive results, consistency/robustness tradeoffs, and experimental results
- Implied results



- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





- Given a set of intervals, select a subset of non-overlapping intervals with maximum cardinality.
- In the offline setting, a simple greedy algorithm solves the problem optimally.





\_\_\_\_











































• An algorithm has a competitive ratio of r iff

 $\forall l : Alg(l) \ge rOpt(l) - o(Opt(l))$ 





• An algorithm has a competitive ratio of r iff

 $\forall l : ALG(l) \ge rOpt(l) - o(Opt(l))$ 

 The best competitive ratio is Θ(1/m) and Θ(1/log m) for deterministic and online algorithms, respectively, where m is the maximum interval length [Awerbuch et al., SODA'94, Lipton & Tomkins, SODA'94].



- Instead of intervals, the input is formed by pairs of vertices in a given graph.
  - The goal is to accept a maximum number of pairs with disjoint paths between them.





- Instead of intervals, the input is formed by pairs of vertices in a given graph.
  - The goal is to accept a maximum number of pairs with disjoint paths between them.





- Instead of intervals, the input is formed by pairs of vertices in a given graph.
  - The goal is to accept a maximum number of pairs with disjoint paths between them.





- Instead of intervals, the input is formed by pairs of vertices in a given graph.
  - The goal is to accept a maximum number of pairs with disjoint paths between them.





## Disjoing Path Allocation Problem

- Instead of intervals, the input is formed by pairs of vertices in a given graph.
  - The goal is to accept a maximum number of pairs with disjoint paths between them.
  - The problem is NP-hard for general graphs (even SP-graphs) [Even and Etai, 1976] and polynomial-time solvable for trees and outerplanar graphs [Garg, Vazirani, and Yannakakis, 1977], Wagner, 1995].





• Online ALgorithms: give worst-case guarantees but do not provide any insight into typical (average-case) performance.



# Conline Algorithms with Prediction

- **Online ALgorithms:** give worst-case guarantees but do not provide any insight into typical (average-case) performance.
- Machine Learning: works well on typical inputs but can go terribly wrong on unusual (worst-case) inputs.



## Online Algorithms with Prediction

- Online ALgorithms: give worst-case guarantees but do not provide any insight into typical (average-case) performance.
- Machine Learning: works well on typical inputs but can go terribly wrong on unusual (worst-case) inputs.
- Online Algorithms with Prediction: get the best of two worlds via potentially erroneous prediction about the input.



### Online Algorithms with Prediction

- Online ALgorithms: give worst-case guarantees but do not provide any insight into typical (average-case) performance.
- Machine Learning: works well on typical inputs but can go terribly wrong on unusual (worst-case) inputs.
- Online Algorithms with Prediction: get the best of two worlds via potentially erroneous prediction about the input.





- What prediction should be?
- How to measure error?
- Algorithm Design and Analysis



• We consider predictions that concern membership in the input sequence.




• We consider predictions that concern membership in the input sequence.





- We consider predictions that concern membership in the input sequence.
- Statistical predictions such as average input length are unlikely to help.





# • We use $\eta = Opt(FP \cup FN)$





- We use  $\eta = Opt(FP \cup FN)$
- Desirable Properties:



- We use  $\eta = \operatorname{Opt}(\mathit{FP} \cup \mathit{FN})$
- Desirable Properties:
  - Monotonicity: eliminating false negatives/positives must not increase the error.



- We use  $\eta = \operatorname{Opt}(\mathit{FP} \cup \mathit{FN})$
- Desirable Properties:
  - Monotonicity: eliminating false negatives/positives must not increase the error.
  - Lipschitz: the error is not "too small".
  - Completeness: the error is not "too large".



- We use  $\eta = \operatorname{Opt}(\mathit{FP} \cup \mathit{FN})$
- Desirable Properties:
  - Monotonicity: eliminating false negatives/positives must not increase the error.
  - Lipschitz: the error is not "too small".
  - Completeness: the error is not "too large".
- We define normalized error  $\gamma(\hat{I}, I) = \frac{\eta(\hat{I}, I)}{OPT(\hat{I}, I)}$



• Design an algorithm that is consistent, robust, and smooth.





• The most obvious algorithm to try (first) just follows the prediction:

Algorithm Trust
From $\hat{I}$ , compute an optimal solution, $I^*$
for all requests $r \in I$ :
if $r \in I^*$ :
accept
else:
reject



• A positive result for general graphs:

# Theorem

On any graph, 
$$\operatorname{TRUST}(\hat{I}, I) \geq (1 - 2\gamma(\hat{I}, I)) \operatorname{OPT}(I)$$

# Proof.

$$\begin{split} \text{TRUST}(\hat{I}, I) &\geq \text{OPT}(\hat{I}) - \text{OPT}(FP) \\ &\geq \text{OPT}(I) - \text{OPT}(FN) - \text{OPT}(FP) \\ &\geq \text{OPT}(I) - 2\text{OPT}(FP \cup FN) \\ &= \text{OPT}(I) - 2\eta(\hat{I}, I) \\ &= (1 - 2\gamma(\hat{I}, I))\text{OPT}(I) \end{split}$$



# Theorem

On a star graphs  $S_8$ ,  $ALG(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) OPT(I)$ 



# Theorem

On a star graphs  $S_8$ ,  $ALG(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) \text{ OPT}(I)$ 

• Showing, equivalently,  $ALG(\hat{I}, I) \leq OPT(I) - 2\eta(\hat{I}, I)$ 



# Theorem

On a star graphs  $S_8$ ,  $ALG(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) OPT(I)$ 

- Showing, equivalently,  $ALG(\hat{I}, I) \leq OPT(I) 2\eta(\hat{I}, I)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.



#### Theorem

On a star graphs  $S_8$ ,  $ALG(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) \text{ Opt}(I)$ 

- Showing, equivalently,  $ALG(\hat{I}, I) \leq OPT(I) 2\eta(\hat{I}, I)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{I} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$





On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$





On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$





On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$





On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$





#### Theorem

On a star graphs  $S_8$ ,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



$$\mathsf{I}=(2,3)$$
 ,  $(6,7)$  ,  $(1,2)$  ,  $(3,4)$ 



On a star graphs 
$$S_8$$
,  $\operatorname{ALG}(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) \operatorname{Opt}(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



$$\mathsf{I}=(2,3)$$
 ,  $(6,7)$  ,  $(1,2)$  ,  $(3,4)$  ,  $(7,8)$ 



#### Theorem

On a star graphs  $S_8$ ,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



$$\mathsf{I}=(2,3)$$
 ,  $(6,7)$  ,  $(1,2)$  ,  $(3,4)$  ,  $(7,8)$ 



#### Theorem

On a star graphs  $S_8$ ,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



 $\mathsf{I}=(2,3)$  , (6,7) , (1,2) , (3,4) , (7,8) , (5,8)



## Theorem

On a star graphs 
$$S_8$$
,  $Alg(\hat{I}, I) \leq (1 - 2\gamma(\hat{I}, I)) Opt(I)$ 

- Showing, equivalently,  $Alg(\hat{l}, l) \leq Opt(l) 2\eta(\hat{l}, l)$
- Exhibit an input (and prediction) s.t., the prediction error is 1, and the profit of OPT is 2 more than the profit of ALG.
- One case of the proof:  $\hat{l} = \{(1,2), (2,3), (3,4), (4,5), (6,7), (7,8)\}$



I = (2,3) , (6,7) , (1,2) , (3,4) , (7,8) , (5,8)  $Opt(FN \cup FP) = Opt(\{(4,5),(5,8)\}) = 1$ 



• As long as a graph class contains  $S_8$ , TRUST (follow-the-prediction) is the best possible.



- As long as a graph class contains  $S_8$ , TRUST (follow-the-prediction) is the best possible.
- This motivates us to focus on interval graphs (interval scheduling).



```
Algorithm TrustGreedy

From Î, compute a left-most optimal solution, I*

for all requests r ∈ I:

if r does not overlap an accepted request and

(is in I* or

does not overlap any I*-requests or

overlaps exactly one I*-request ending no earlier than r)

accept r

update I* if necessary

else:

reject r
```




































Online Interval Scheduling with Predictions



































For any prediction  $\hat{l}$  and input sequence l, TRUSTGREEDY $(\hat{l}, l) \ge (1 - \gamma(\hat{l}, l))$  OPT(l)

• An improvement over the competitive ratio  $1 - 2\gamma(\hat{I}, I)$  of TRUST.





For any deterministic algorithm ALG, there are input sequences and predictions I and  $\hat{I}$ , so ALG $(\hat{I}, I) \leq (1 - \gamma(\hat{I}, I))$  OPT(I)

• Let  $\hat{I} = \{(0,2), (0,1)\}$ , and I start with (0,2).



- Let  $\hat{I} = \{(0,2), (0,1)\}$ , and I start with (0,2).
  - If ALG rejects (0,2), OPT accepts it and input ends, so  $(0,1) \in FP$ , and  $\eta = 1$ .



- Let  $\hat{I} = \{(0,2), (0,1)\}$ , and I start with (0,2).
  - If ALG rejects (0,2), OPT accepts it and input ends, so  $(0,1) \in FP$ , and  $\eta = 1$ .
  - If ALG accepts (0,2), then I continues with (0,1) and (1,2) that OPT accepts, so  $(1,2) \in FN$  and  $\eta = 1$ .

$$(0, 2)$$

$$Alg = 0 \quad Opt = 1$$

$$\eta = 1$$



- Let  $\hat{I} = \{(0,2), (0,1)\}$ , and I start with (0,2).
  - If ALG rejects (0,2), OPT accepts it and input ends, so  $(0,1) \in FP$ , and  $\eta = 1$ .
  - If ALG accepts (0,2), then I continues with (0,1) and (1,2) that OPT accepts, so  $(1,2) \in FN$  and  $\eta = 1$ .

(0, 2)	(0, 2)	
Alg = 0  Opt = 1 $\eta = 1$	(0, 1)	(1, 2)
	Alg = 1	Opt=2
7	n = 1	



• **Consistency** refers to the competitive ratio when the predictions are correct, and **robusteness** is the competitive ratio when predictions are adversarial.



- **Consistency** refers to the competitive ratio when the predictions are correct, and **robusteness** is the competitive ratio when predictions are adversarial.
- Starting with a negative result:

If a (possibly randomized) algorithm ALG is both  $\alpha$ -consistent, then its robustness is at most  $\beta = \frac{2(1-\alpha)}{|\log m|-1}$ .



• For a positive result, we define ROBUST-TRUST( $\alpha$ ) as follows:

#### Algorithm RobustTrust ( $\alpha$ )

Draw probablity p uniformly at random if  $p < \alpha$ : apply algorithm TrustGreedy else apply algorithm Classify-and-Randomly-Select



• For a positive result, we define  $ROBUST-TRUST(\alpha)$  as follows:

```
Algorithm RobustTrust (\alpha)

Draw probablity p uniformly at random

if p < \alpha:

apply algorithm TrustGreedy

else

apply algorithm Classify-and-Randomly-Select
```

## Theorem

ROBUST-TRUST( $\alpha$ ) has consistency at least  $\alpha$  and robustness at least  $\frac{1-\alpha}{\lceil \log m \rceil}$ 

• Robust-Trust( $\alpha$ ) asymptotically Pareto optimal.



• Consider TRUST, TRUSTGREEDY, GREEDY, and OPT on real-world scheduling data on parallel machines [Chapin et al. IPPS/SPDP, 1999]





• The negative result on star graphs implies a negative result for **matching** in general graphs (even if restricted to planar graphs).





- The negative result on star graphs implies a negative result for matching in general graphs (even if restricted to planar graphs).
- The negative results on matching implies a negative result for independent set in general graphs



- For disjoint path allocation problem, TRUST has a competitive ratio of  $1 2\gamma(\hat{I}, I)$ , which is optimal.
- For interval scheduling, TRUSTGREEDY has a competitive ratio of  $1 \gamma(\hat{l}, l)$ , which is optimal.
- For consistency/robustness tradeoff, ROBUSTTRUST( $\alpha$ ) is  $\alpha$ -consistent and  $(1 \alpha)/\lceil \log m \rceil$ -robust, which is asymptotically Pareto-optimal.