Learning Thresholds for PV Change Detection from Operators' Labels

Dapeng Liu, Youjian Zhao, Kaixin Sui, Shiwen Cheng, Dan Pei, Chengbin Quan
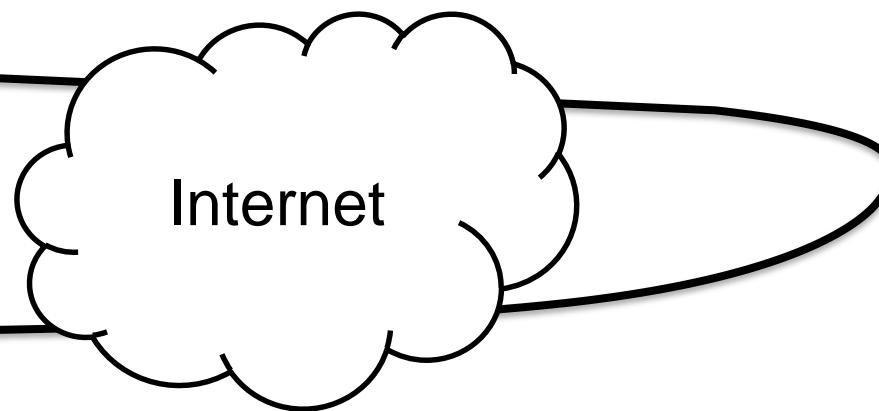Jiao Luo, Xiaowei Jing, Mei Feng

PVs (Page Views)

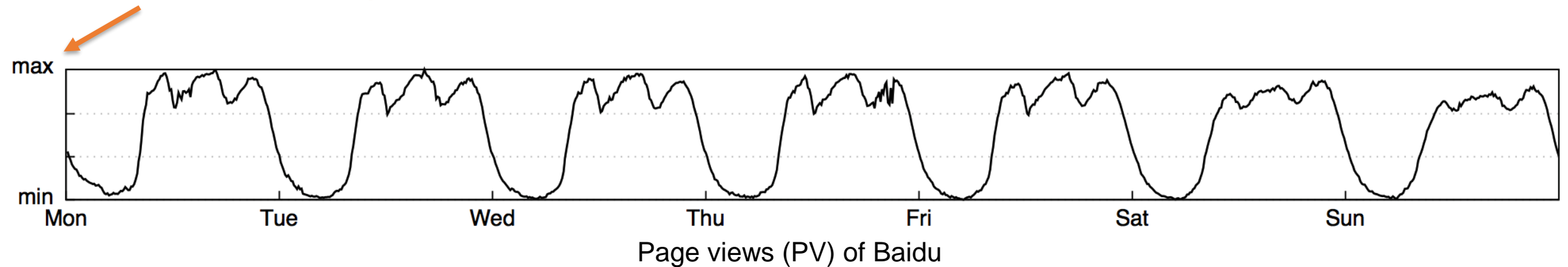When you submit a query and get the result, you contribute a **search PV (Page View)**
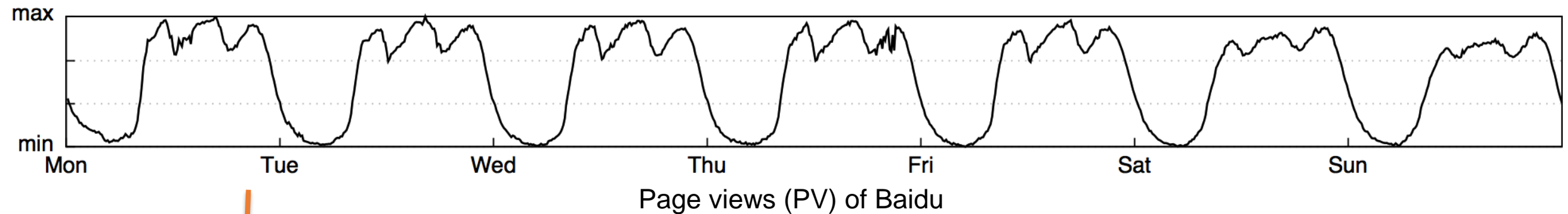
User

Internet

Server

PVs (Page Views)

- **PV** is crucial for search engines (e.g., revenue and market share)

Over a million PVs per minute

Page views (PV) of Baidu

PVs (Page Views)

- **PV** is crucial for search engines (e.g., revenue and market share)
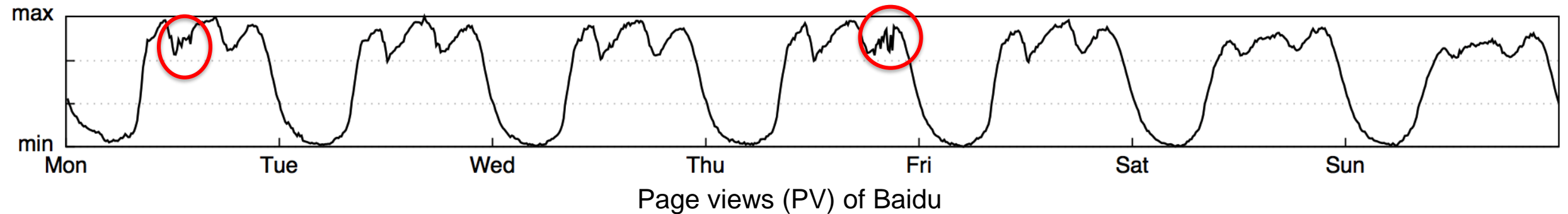

Page views (PV) of Baidu

PV breakdowns
in different ways

- ISPs (e.g., Nanjing Telecom, Beijing Unicom)
- Devices (e.g., mobile phones, PCs)
- Data centers
- …

PVs (Page Views)

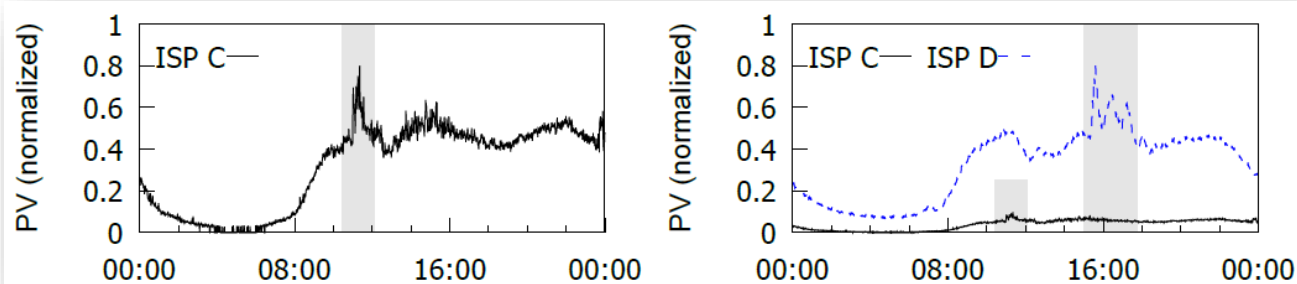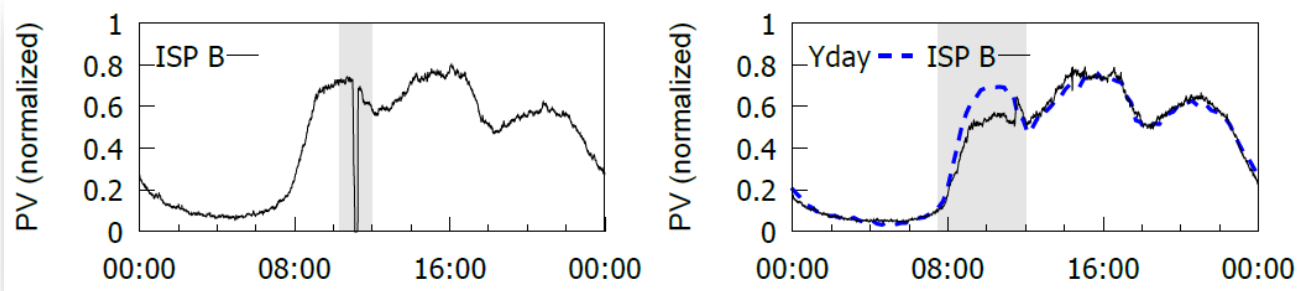- **PV** is crucial for search engines (e.g., revenue and market share)
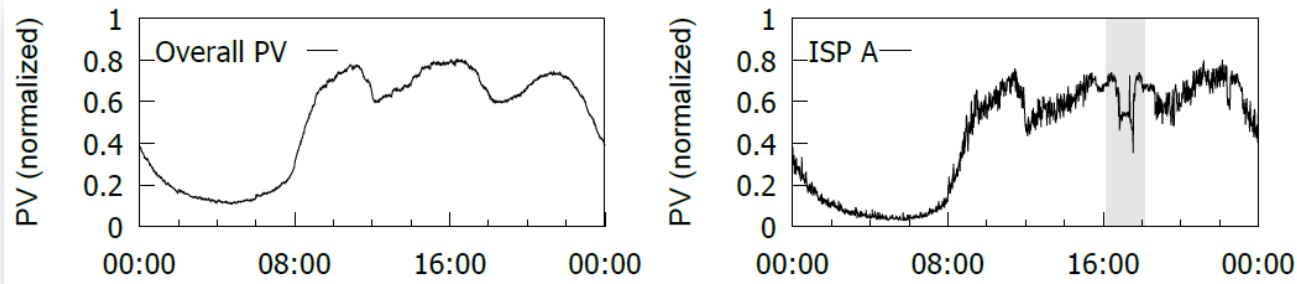

Page views (PV) of Baidu

Significant PV changes → Potential failures, bugs, attacks...

Operators are responsible for detection

What changes do they care?

Significant PV Changes

Overall PV is not efficient
- **Fine-grained PVs from 100+ ISPs**

Change properties
- **Change severity & duration**

ISP PVs are not born equal
- **Large ISP is more important**

Significant PV Changes

Problem

Manfully tune many change severity & duration thresholds for different ISPs

Overall PV is not efficient
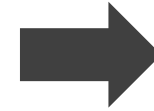- **Fine-grained PVs from 100+ ISPs**

Change properties
- **Change severity & duration**

ISP PVs are not born equal
- **Large ISP is more important**

Significant PV Changes

Problem

Manfully tune many change severity & duration thresholds for different ISPs

➡️

Idea

Automatically tune thresholds from operators' feedback of the detection results

Significant PV Changes

Problem

Manfully tune many change severity & duration thresholds for different ISPs
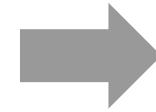
Idea

Automatically tune thresholds from operators' feedback of the detection results

Challenges

- Operators only provide a few labels
- Labels can be inaccurate sometimes

Significant PV Changes

Problem

Manfully tune many change severity & duration thresholds for different ISPs

Idea

Automatically tune thresholds from operators' feedback of the detection results
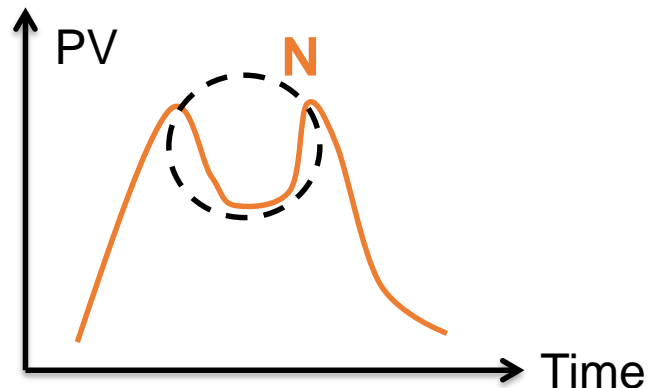
Solution
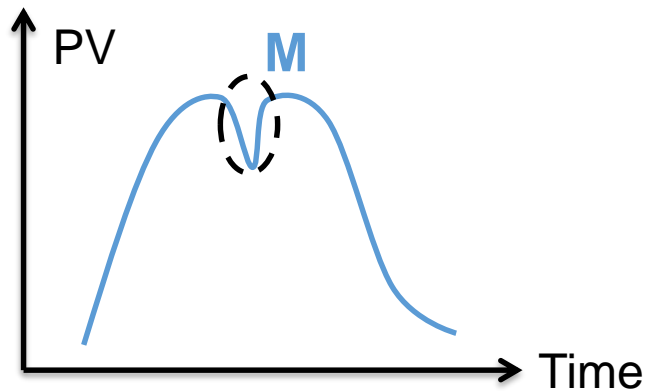
Make use of every label

Challenges

- Operators only provide a few labels
- Labels can be inaccurate sometimes

- Change significance **monotonically increases** with the severity and the duration （for the same ISP）

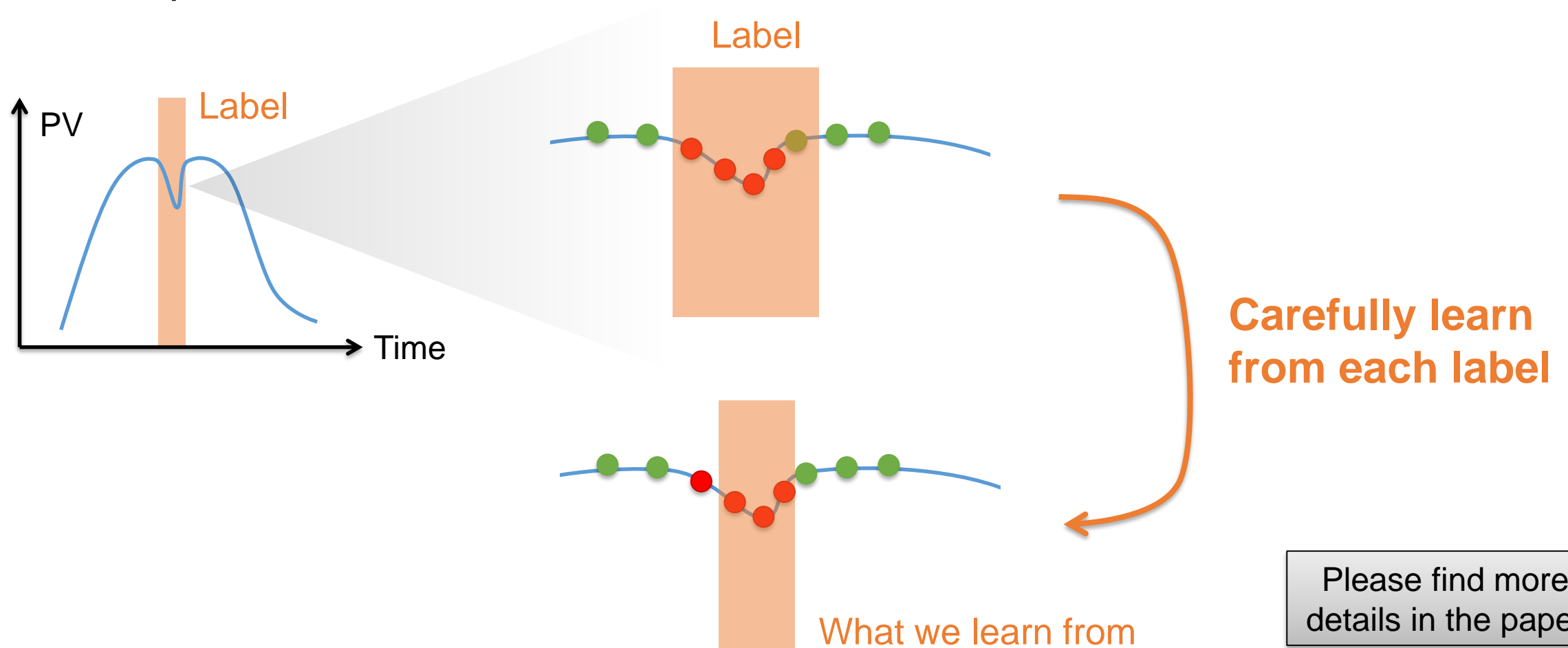If **M** is a significant change, so should be **N**
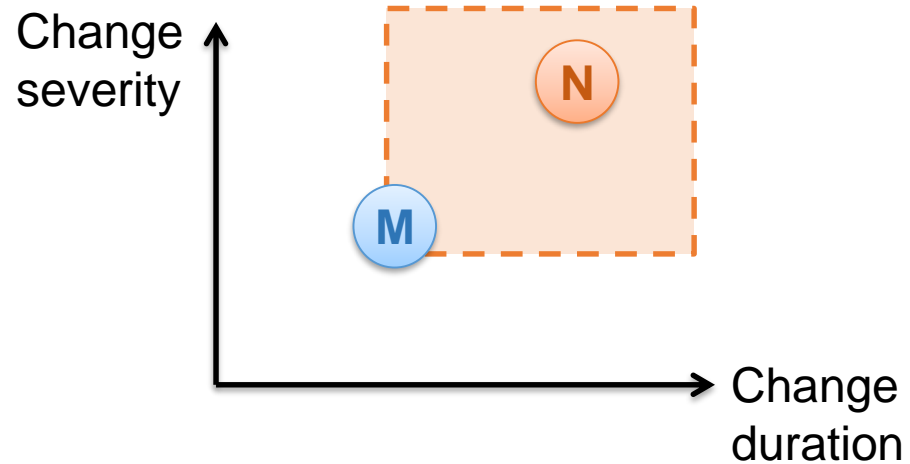
Observation 2

- Most of the label interval is correct, though not 100%
  - Assumption: over a half

Carefully learn from each label

Please find more details in the paper

Observation 3

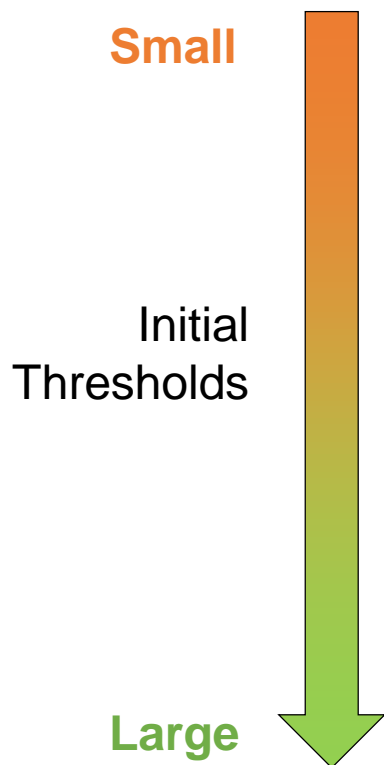- Larger ISPs are more important → **Sharing learning outcomes**
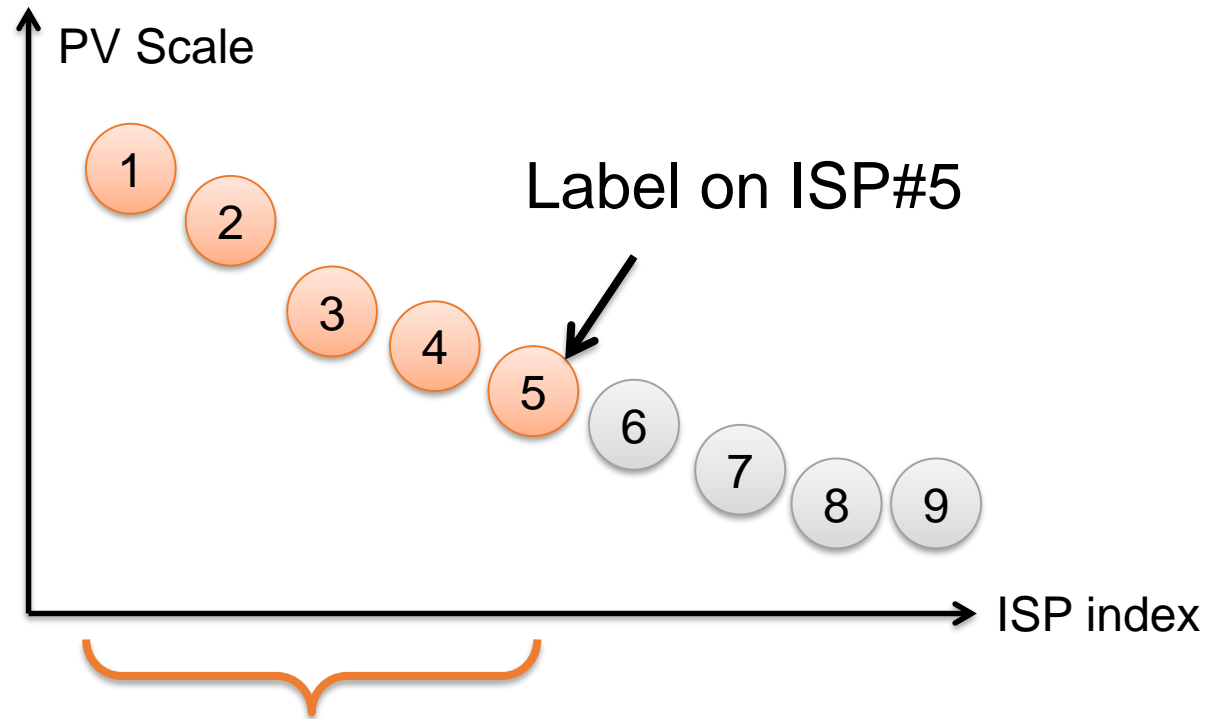
Share the learning result among larger ISPs
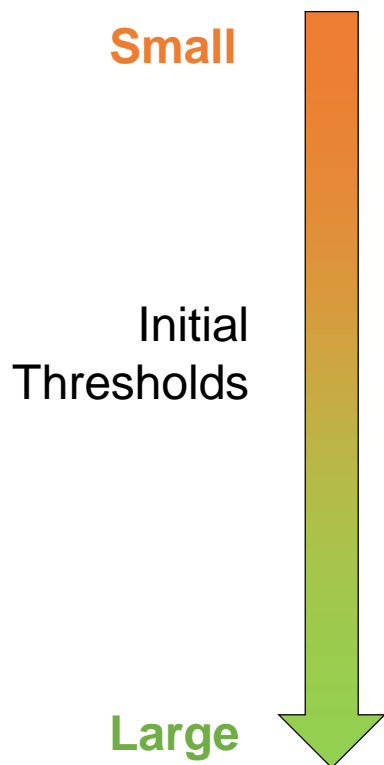
Please find more details in the paper

- Ground truth: real tickets
- 4-month overall PV from Baidu

Small

Initial
Thresholds

Large

Initial $\mathcal{T}$	Threshold Learning	#Alarms	Identified Tickets(%)	#Labels
$\{[2, 1.96]\}$	Fixed Thld	3744	100%	-
	Lazy OP	84	79%	16
	Careful OP	54	87%	25
$\{[3, 3.92]\}$	Fixed Thld	108	83%	-
	Lazy OP	39	83%	4
	Careful OP	51	83%	25
$\{[5, 5.88]\}$	Fixed Thld	18	67%	-
	Lazy OP	18	79%	4
	Careful OP	30	79%	20
$\{[7, 7.84]\}$	Fixed Thld	13	50%	-
	Lazy OP	25	58%	4
	Careful OP	24	75%	9
$\{[10, 9.8]\}$	Fixed Thld	4	17%	-
	Lazy OP	13	46%	2
	Careful OP	34	67%	19

Result # 1

- Ground truth: Real tickets
- 4-month overall PV from Baidu

Small → **Large** (Initial Thresholds)

Initial $\mathcal{T}$	Threshold Learning	#Alarms	Identified Tickets(%)	#Labels
$\{[2, 1.96]\}$	Fixed Thld	3744	100%	-
	Lazy OP	84	79%	16
	Careful OP	54	87%	25
$\{[3, 3.92]\}$	Fixed Thld	108	83%	-
	Lazy OP	39	83%	4
	Careful OP	51	83%	25
$\{[5, 5.88]\}$	Fixed Thld	18	67%	-
	Lazy OP	18	79%	4
	Careful OP	30	79%	20
$\{[7, 7.84]\}$	Fixed Thld	13	50%	-
	Lazy OP	25	58%	4
	Careful OP	24	75%	9
$\{[10, 9.8]\}$	Fixed Thld	4	17%	-
	Lazy OP	13	46%	2
	Careful OP	34	67%	19

Result # 1

- Ground truth: Real tickets
- 4-month overall PV from Baidu

Small → **Large** (Initial Thresholds)

| Initial $\mathcal{T}$	Threshold Learning	#Alarms	Identified Tickets(%)	#Labels
$\{[2, 1.96]\}$	Fixed Thld	3744	100%	-
	Lazy OP	84	79%	16
	Careful OP	54	87%	25
$\{[3, 3.92]\}$	Fixed Thld	108	83%	-
	Lazy OP	39	83%	4
	Careful OP	51	83%	25
$\{[5, 5.88]\}$	Fixed Thld	18	67%	-
	Lazy OP	18	79%	4
	Careful OP	30	79%	20
$\{[7, 7.84]\}$	Fixed Thld	13	50%	-
	Lazy OP	25	58%	4
	Careful OP	24	75%	9
$\{[10, 9.8]\}$	Fixed Thld	4	17%	-
	Lazy OP	13	46%	2
	Careful OP	34	67%	19

- Ground truth: pair-wise validation
- 4-month 103 ISP PV from Baidu

Large initial thresholds

- Ground truth: pair-wise validation
- 4-month 103 ISP PV from Baidu

Small initial thresholds

Conclusion
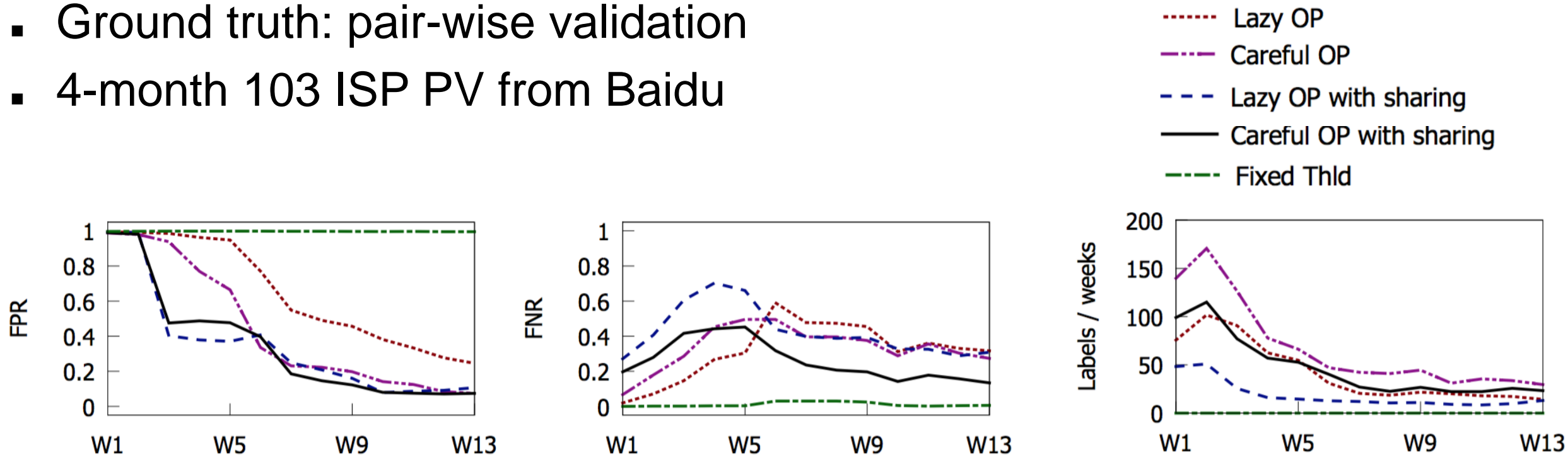
- Automatically tune hundreds of thresholds from labels

- Learn the intention behind each label as much as possible

- Increase the accuracy significantly with tens of labels a week

Thank you