Vadim Denisov, Dirk Fahland, Wil van der Aalst

Predictive Performance Monitoring of Material Handling Systems Using the Performance Spectrum

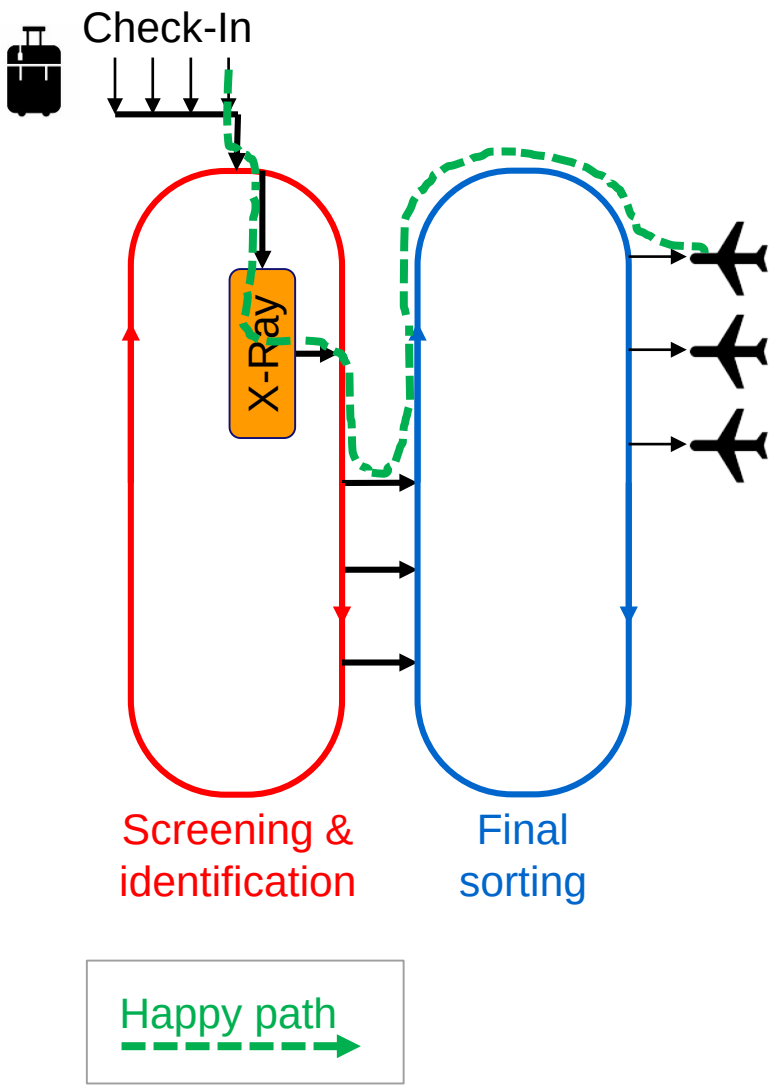
VANDERLANDE

RWTH AACHEN UNIVERSITY

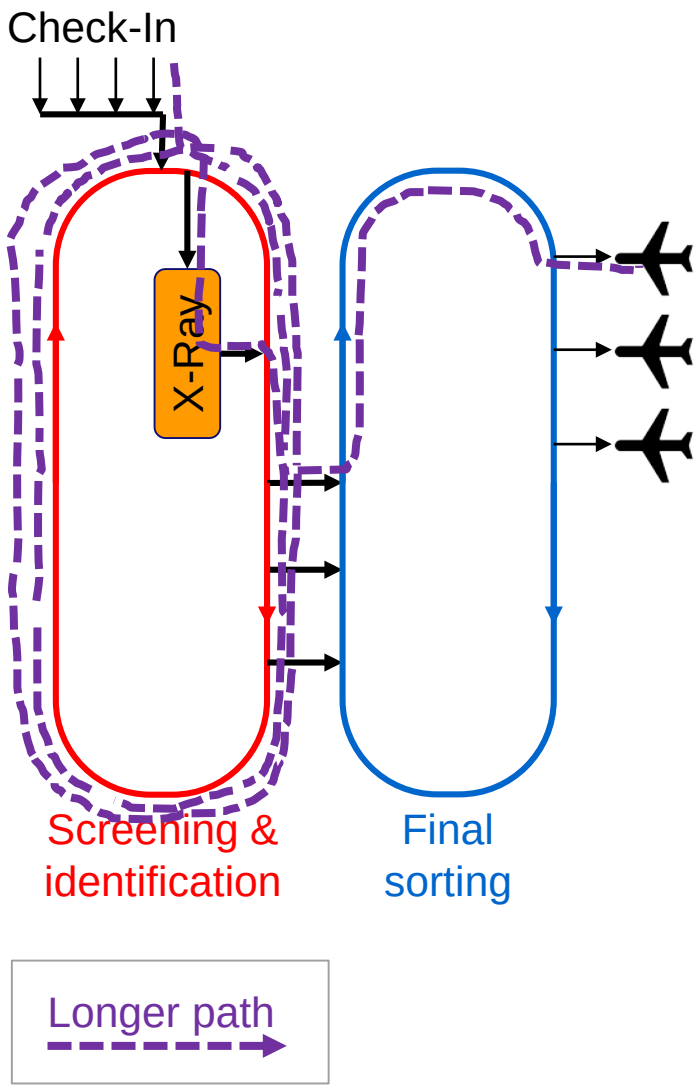
TU/e Technische Universiteit Eindhoven University of Technology

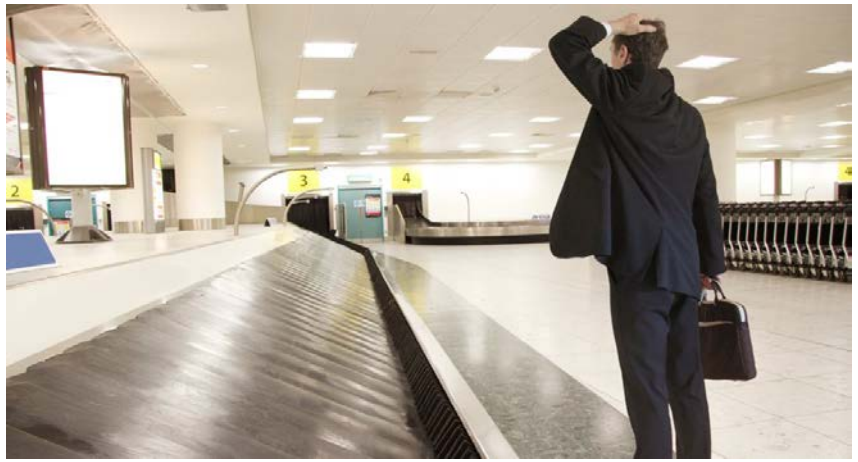
Where innovation starts

Baggage Handling Systems

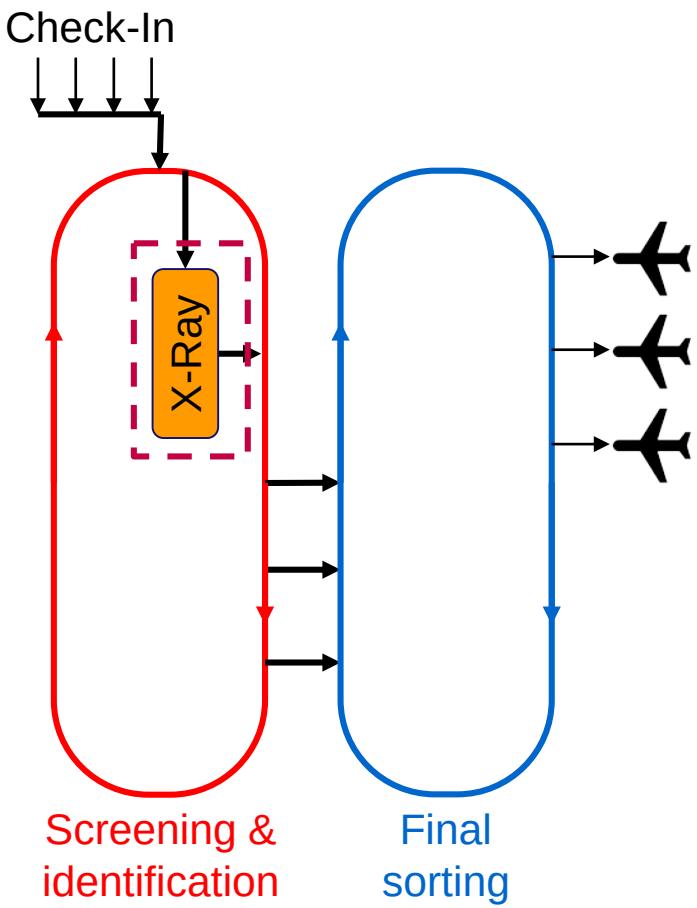


Re-Circulation Causes Significant Delays

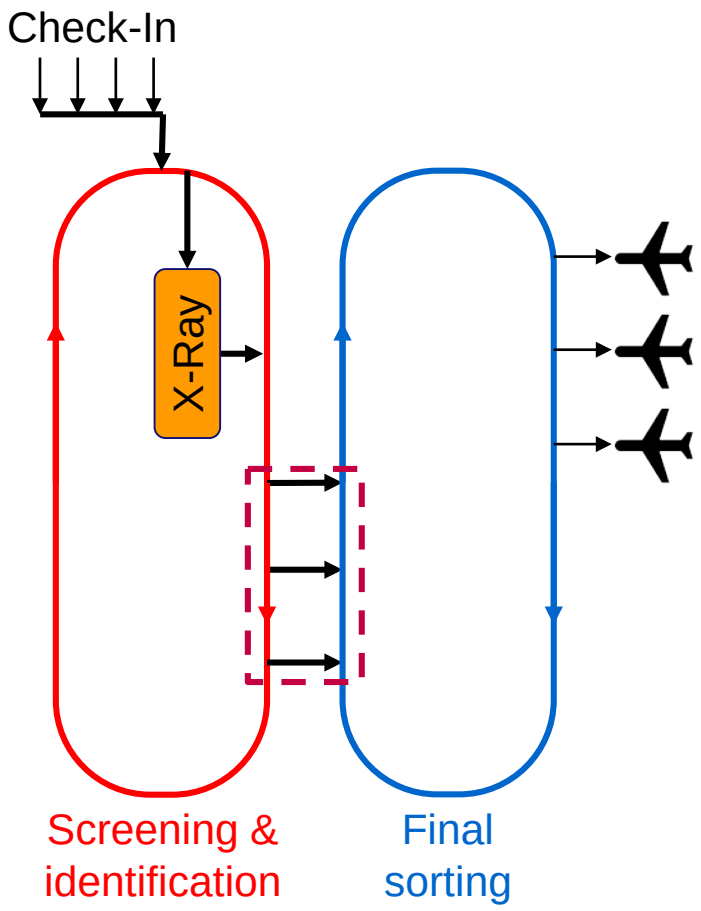




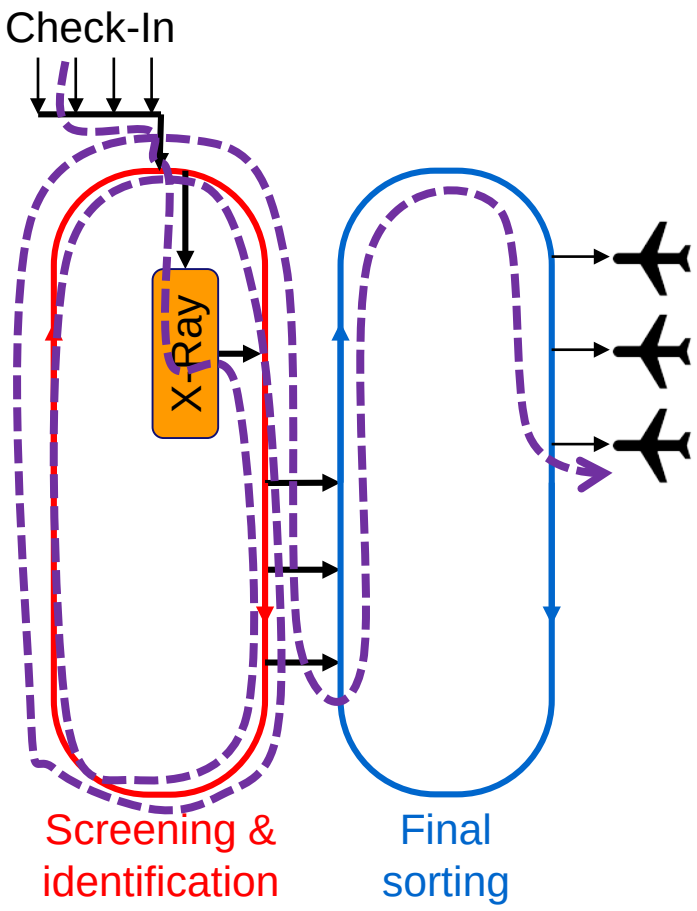
Insufficient Number of Screening Machine Operators



Unavailability of Outgoing Conveyors

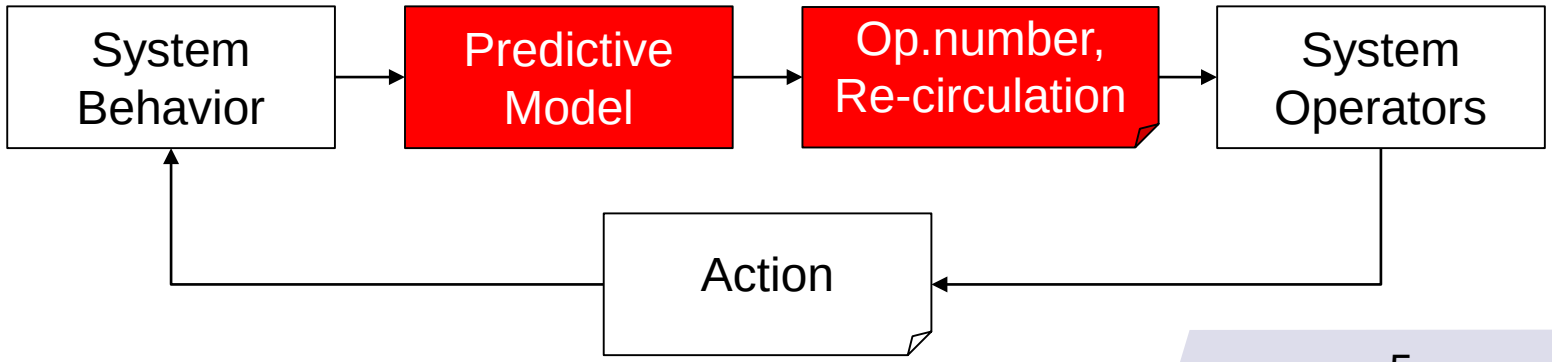


Predictive Performance Monitoring of Material Handling Processes



What is the required **number of operators** in +x minutes from now?

What is the **amount of re-circulation** in +x minutes from now?



What Are Relevant Features?

What is the **load on the screening machines**?

What is the **amount of re-circulation** in +x minutes from now?

One bag = one case

Intra-case features

Bag properties

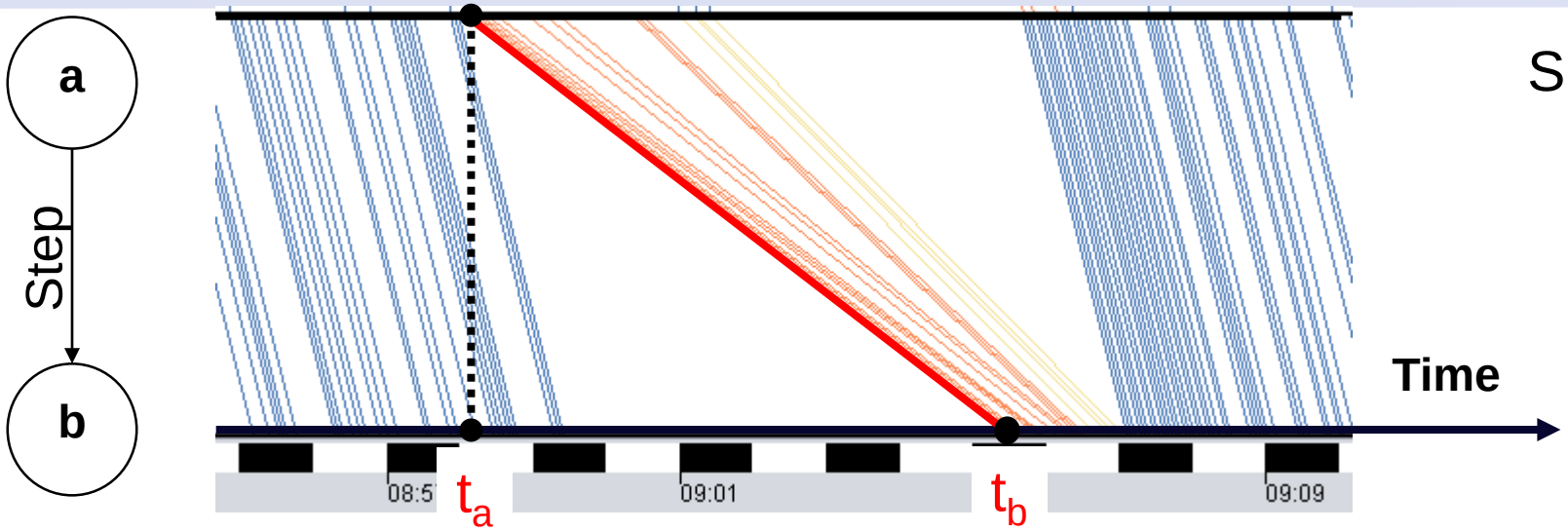
- Bag color: *red*
- Final destination: *Amsterdam*
- Dimensions: *550x450x250mm*
- Weight: *15kg*

Inter-case features

System properties

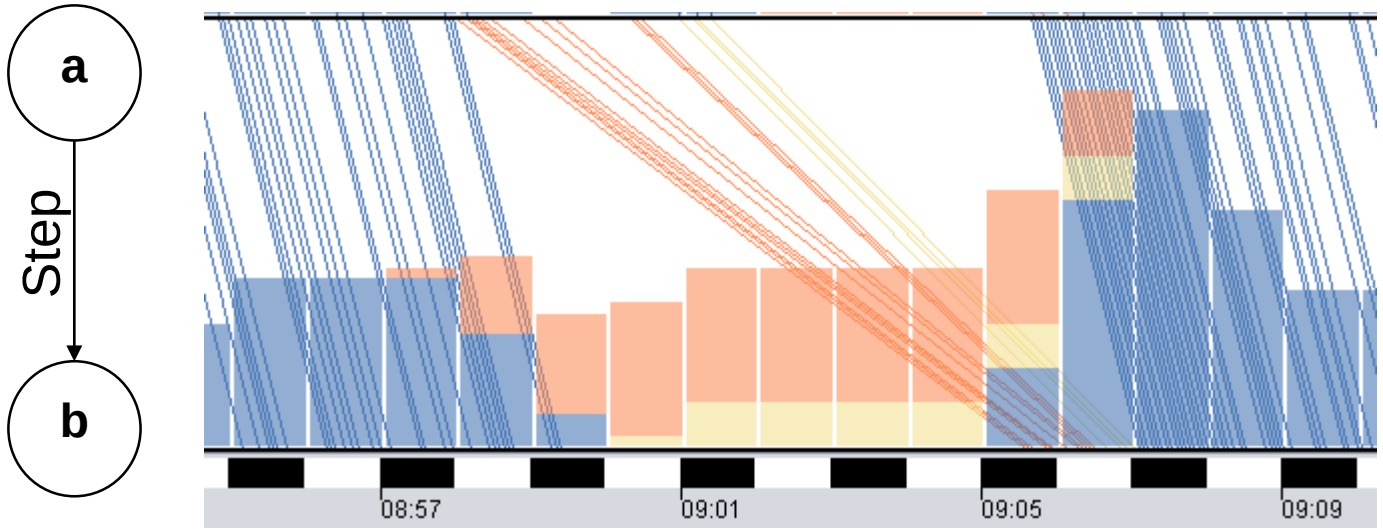
- Utilization of the Loop Sorter: *90%*
- Utilization of conveyors around the Sorter: *60%*
- Current load on the X-Ray Screening Machine: *15 bags/min.*
- Number of bags in the system: *8000*

Performance Spectrum [1]



Sub-trace: $(e_i^{(a,\ ta)}, e_{i+1}^{(b,\ tb)})$

normal speed | 2 times slower | 3 times slower | very slow

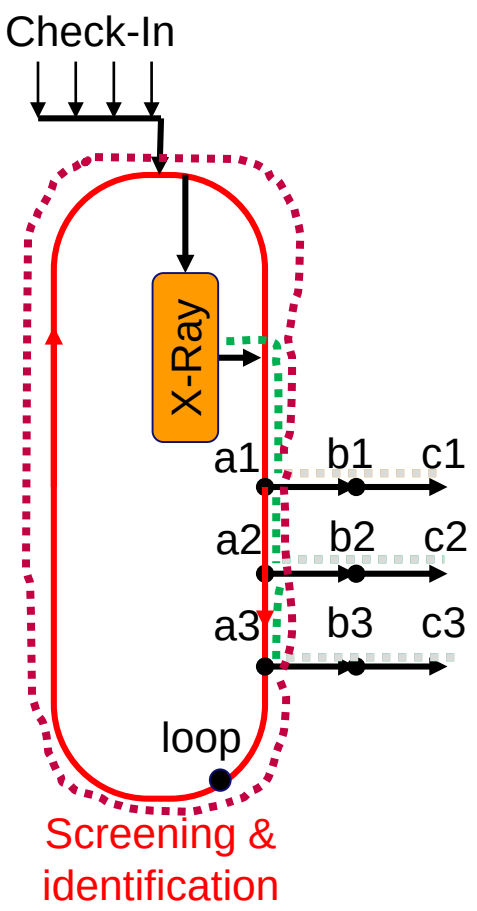


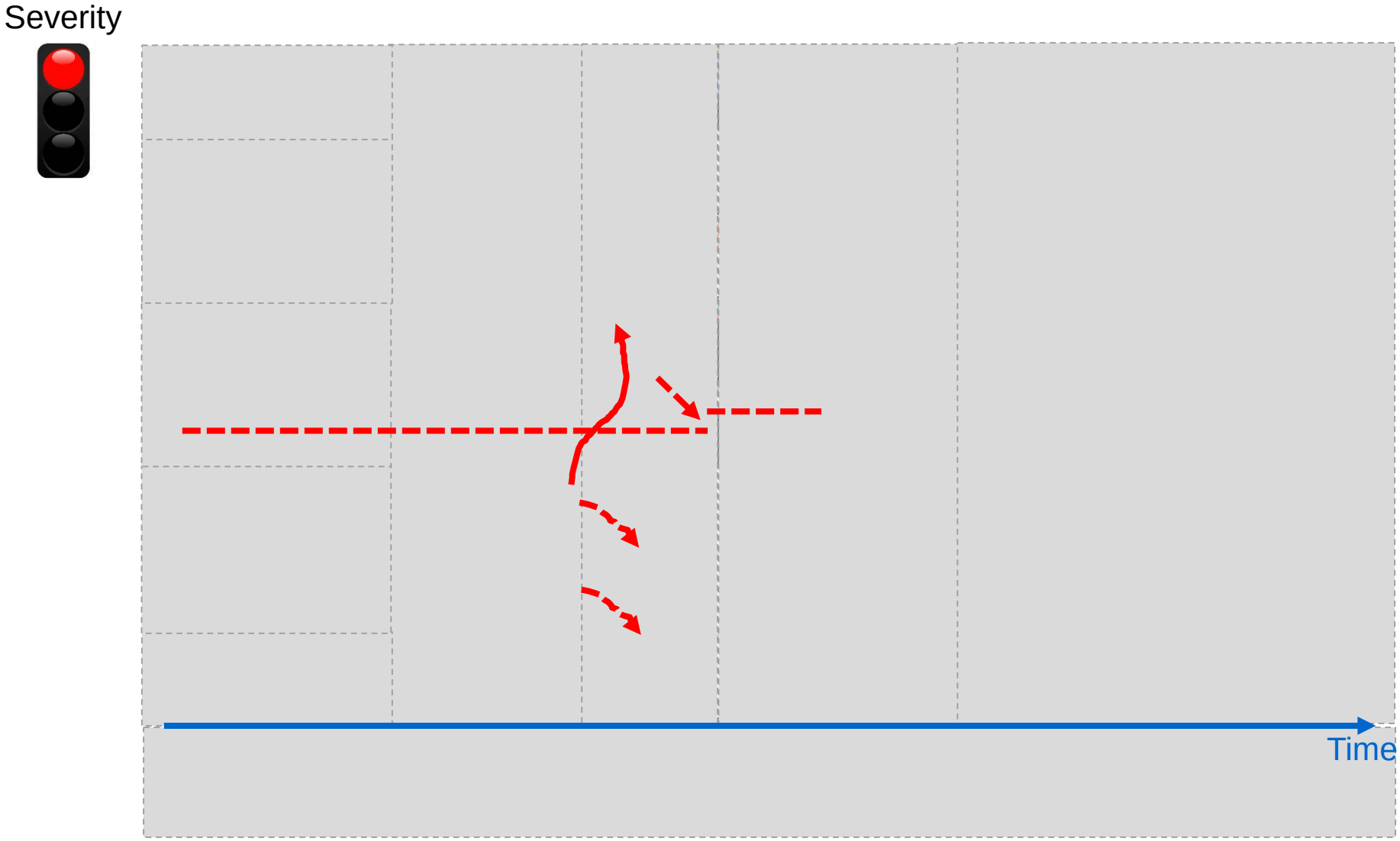
Bins:

- cases **started**
- cases **stopped**
- cases **pending**

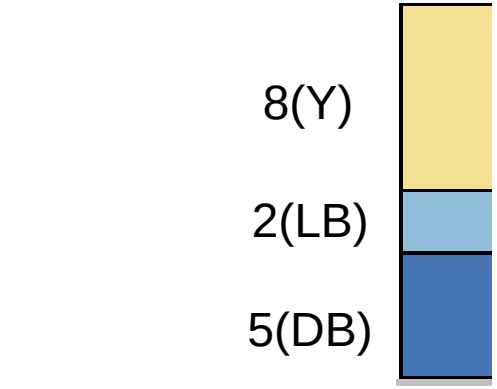
[1] V. Denisov, D. Fahland, and W. M. P. van der Aalst, “Unbiased, finegrained description of processes performance from event data,” in *Business Process Management*, M. Weske, M. Montali, I. Weber, and J. vom Brocke, Eds. Cham: Springer International Publishing, 2018, pp. 139–157.

Development of Problem Instance 2: Re-Circulation because of Unavailability of a1-a3

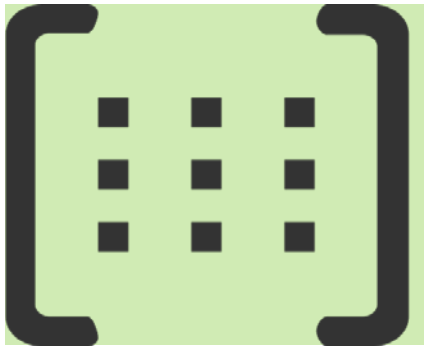


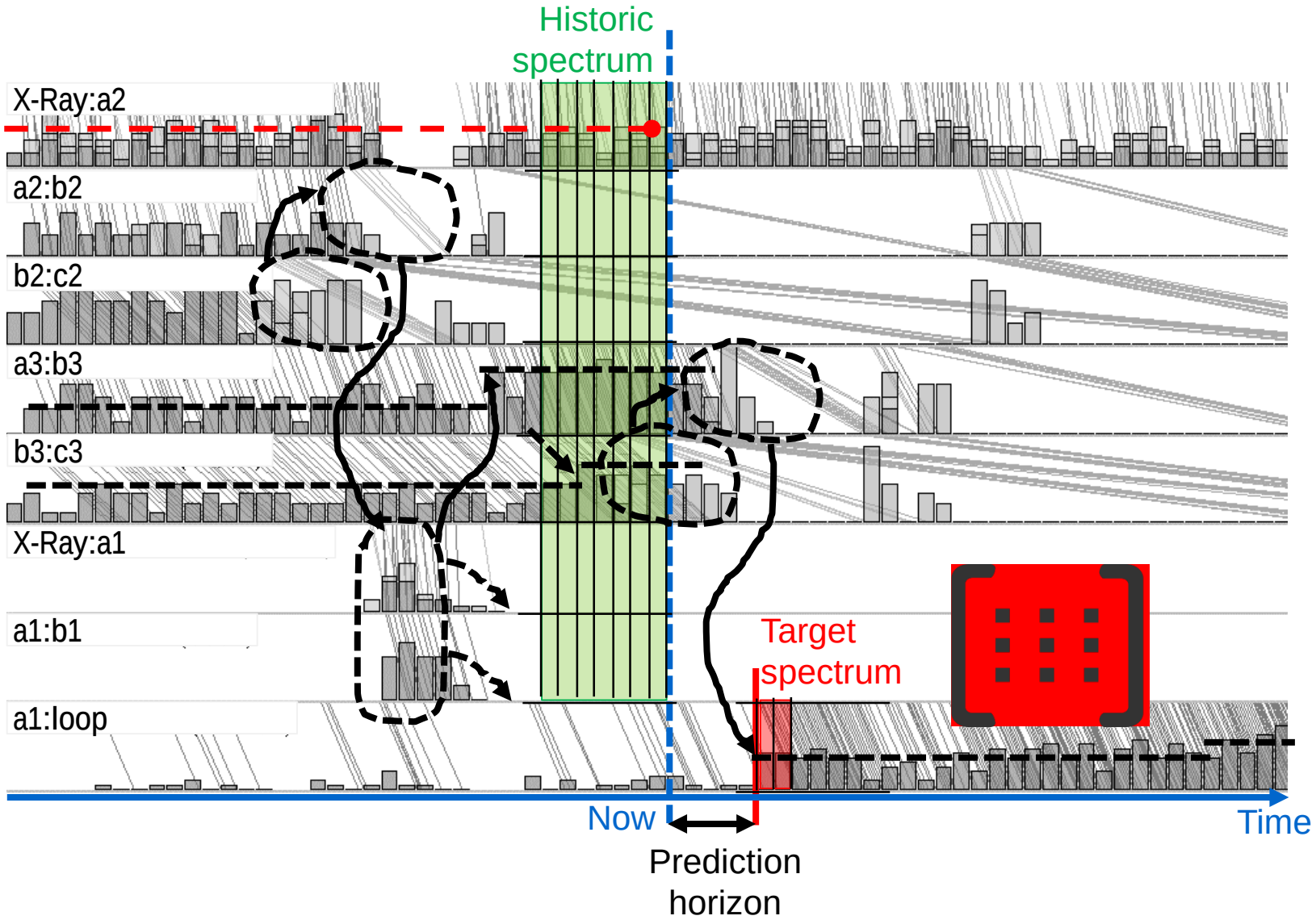


Extracting Features from Performance Spectrum



v_i^j= [8, 2, 5]



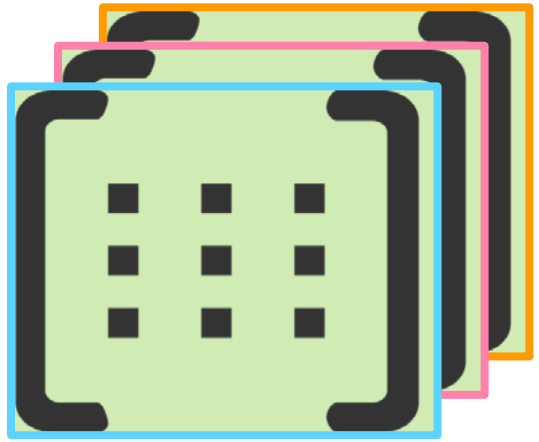


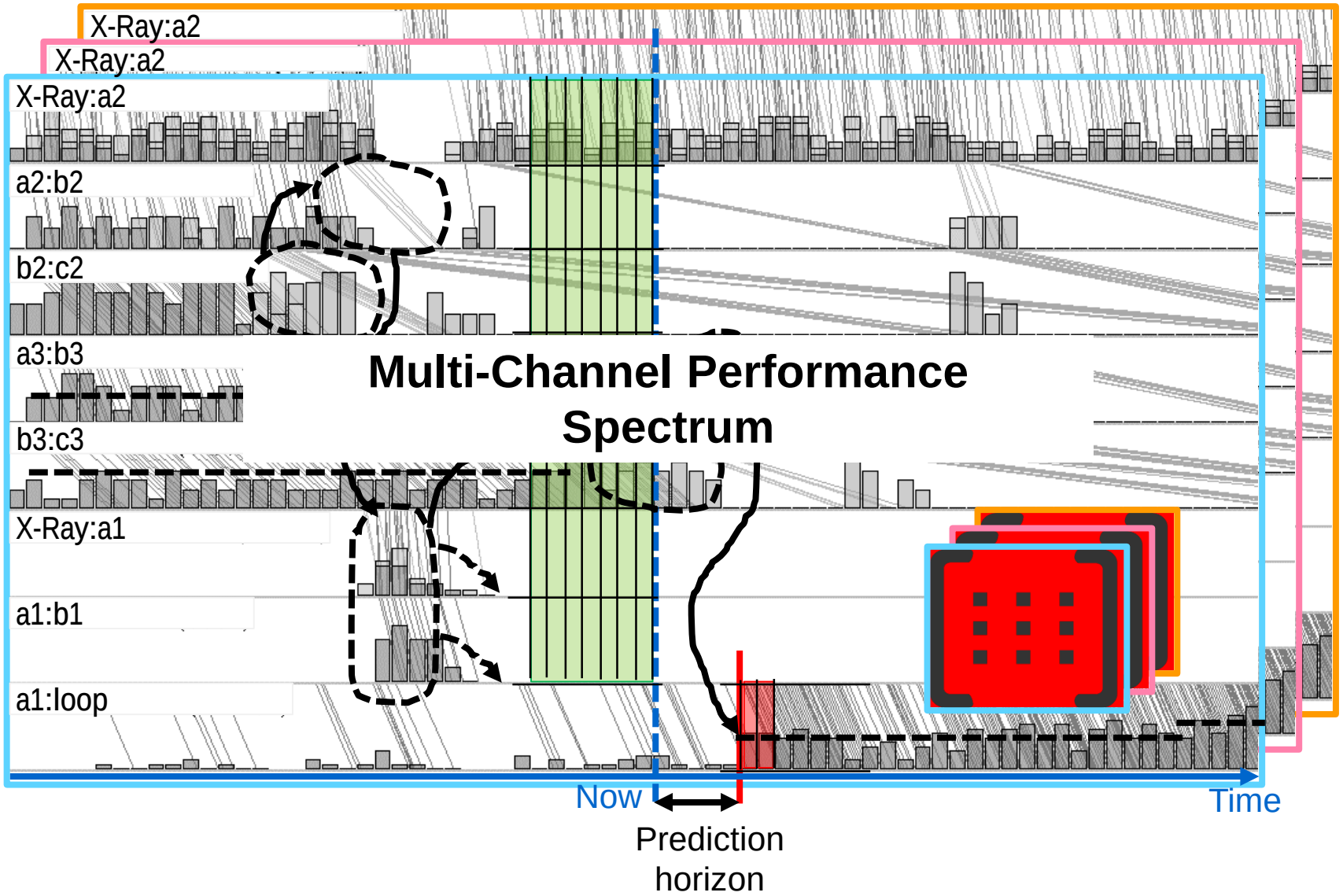
Multi-Channel Performance Spectrum

Channel 1: Speed

Channel 2: Screening result

Channel 3: Sortation status





Multi-Channel Performance Spectrum

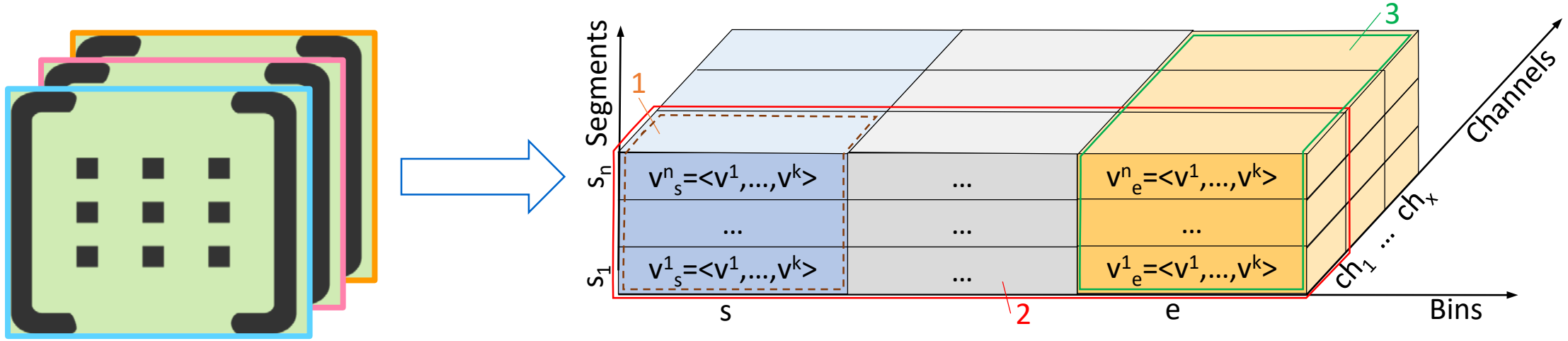
Channel 1: Speed

Channel 2: Screening result

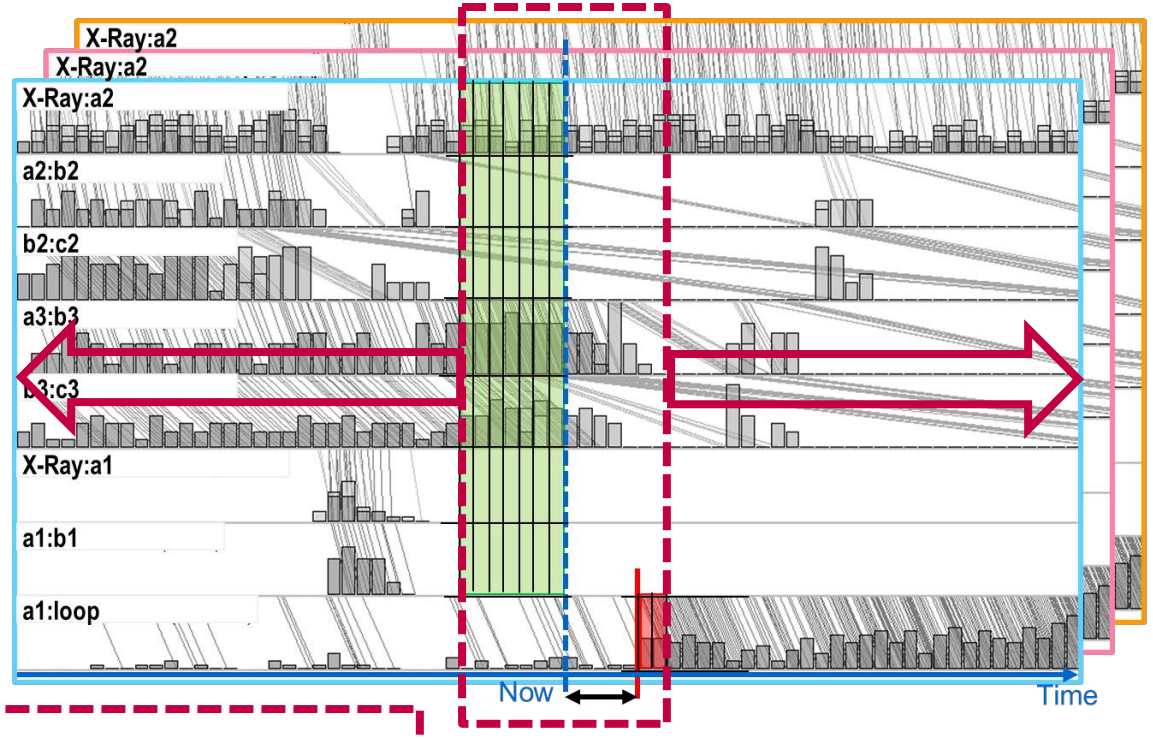
Channel 3: Sortation status

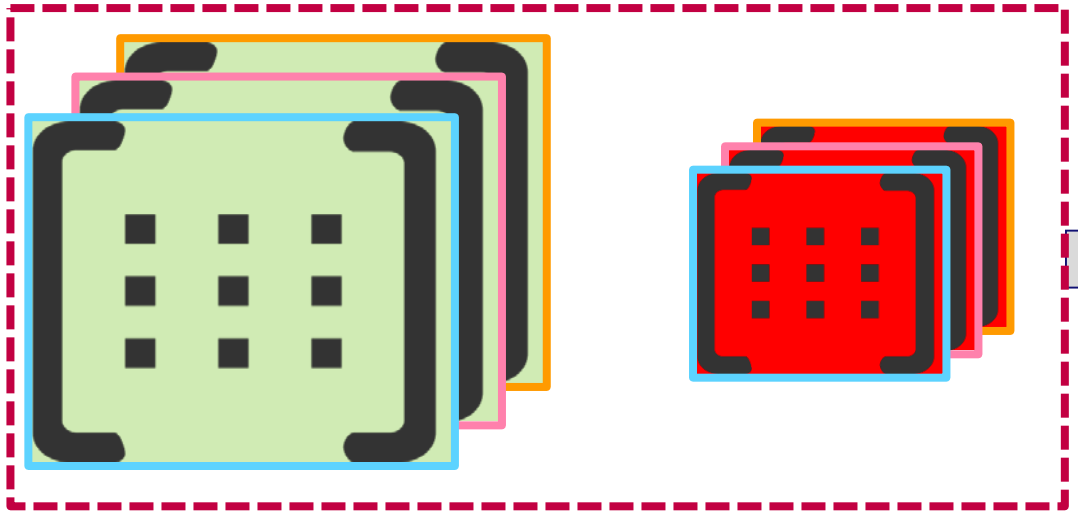
In the paper:

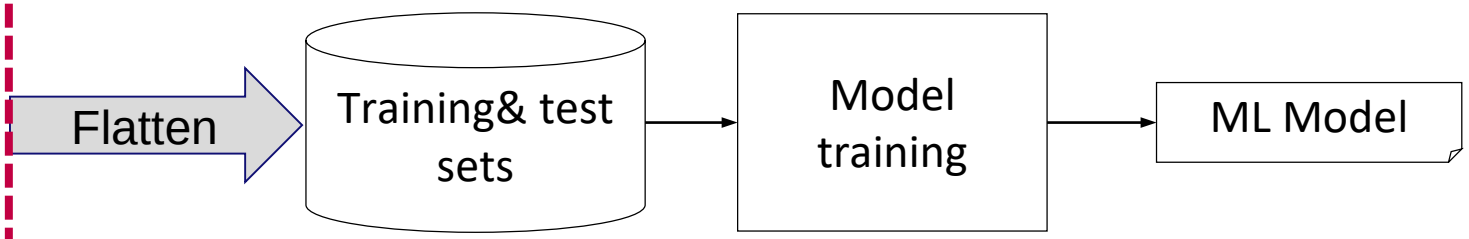
- ✓ **Generic definition** of the **Multi-Channel Performance Spectrum**
- ✓ **Generic method** for **Performance Spectrum -based feature extraction**



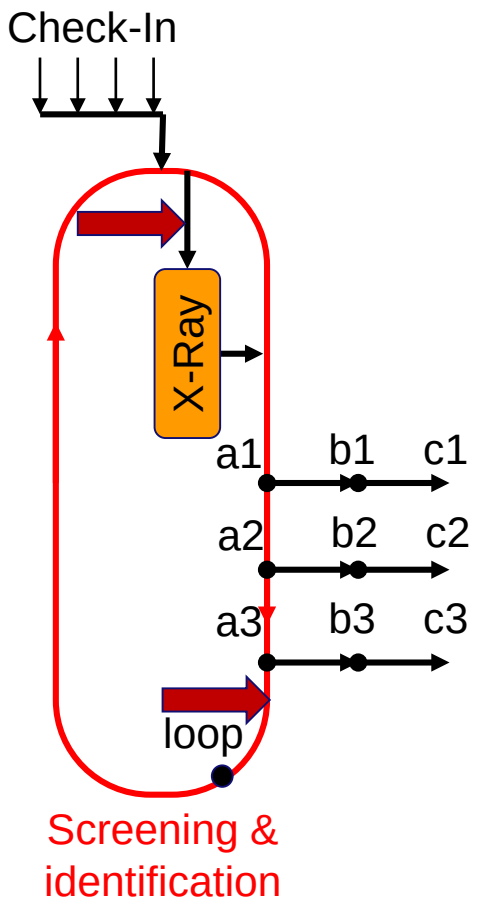
Sliding Window over Performance Spectrum







Problem Instances and Datasets



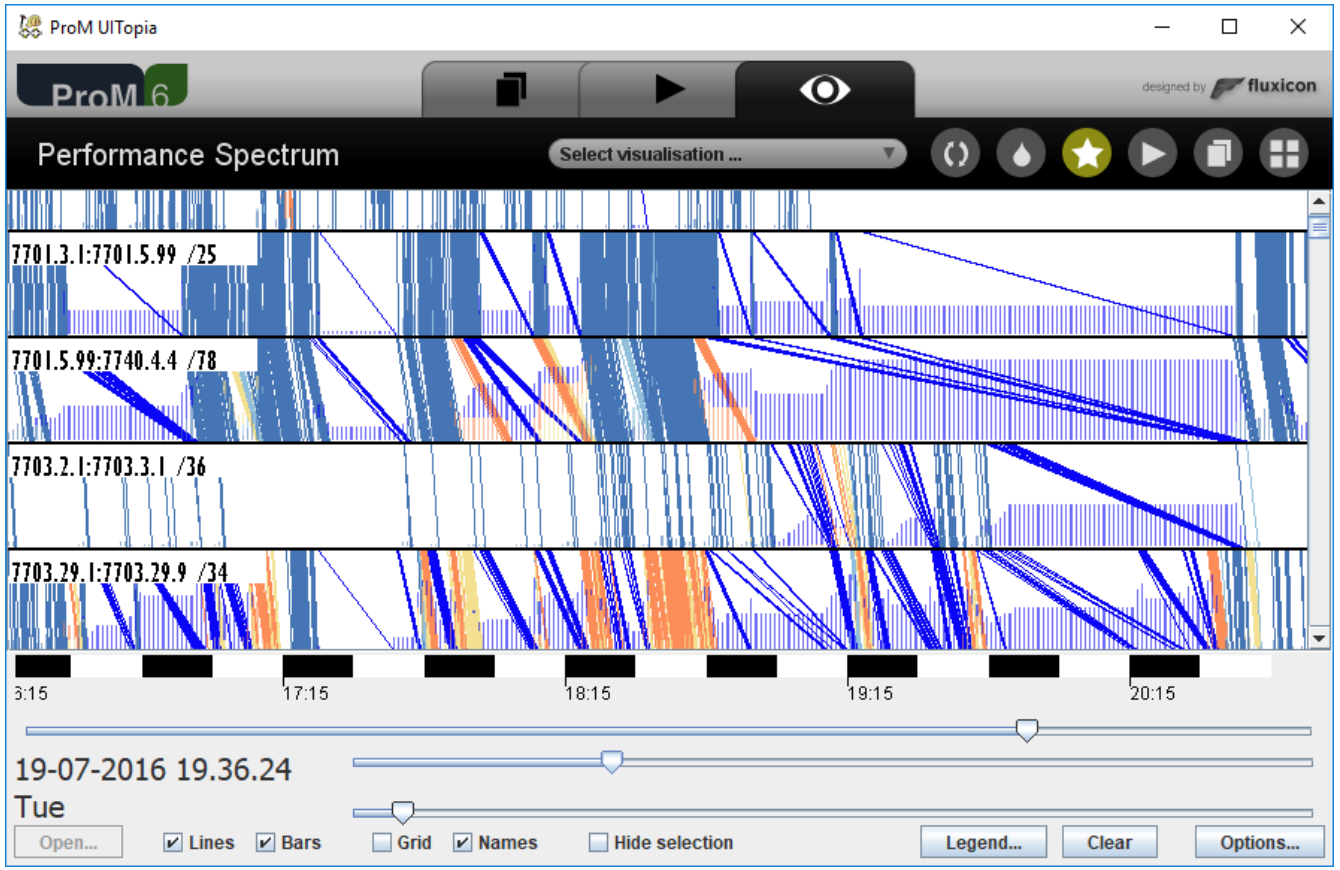
Problem Instance 1
Load on the X-Ray Screening Machines

Problem Instance 2
Extra re-circulation because of unavailability a1-a3

Property	Simulation model [2]	Real BHS of a major European airport
Time period	7 days	4 month
Activities	25	850
Cases per day	1600	25.000-50.000
Events per day	19.000	1.000.000-2.000.000
Events, total	134.000	148.000.000

[2] https://github.com/processmining-in-logistics/psm/releases/tag/1.1.6

Implementation: Performance Spectrum Miner [3] and Scripts on Top of PyTorch



Computational environment:

- a single-machine configuration
- 40 CPUs
- 6 GPUs
- 400GB RAM
- the Performance Spectrum Miner in ProM 6.9
- the PyTorch Framework

http://www.promtools.org/doku.php?id=prom69

https://github.com/processmining-in-logistics/psm/tree/ppm

[3] Denisov, V., Belkina, E., Fahland, D., van der Aalst, W.M.P.: The performance spectrum miner: Visual analytics for fine-grained performance analysis of processes. In: *BPM 2018 Demos.* CEUR Workshop Proceedings, vol. 2196, pp. 96–100. CEUR-WS.org (2018)

Models and Baselines

Models:

- Logistic Regression (LR)
- Feedforward Neural Network (FFNN)

Baselines:

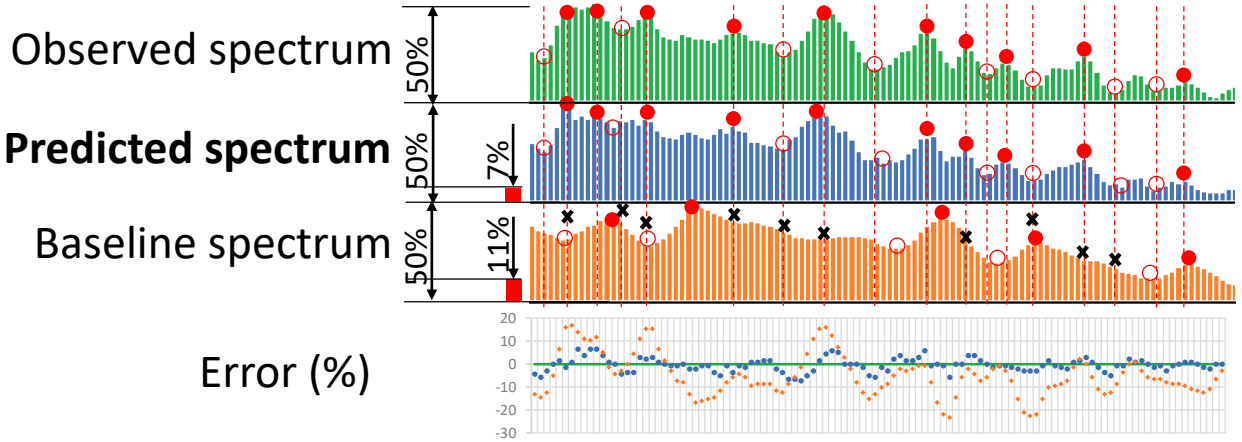
- Naïve baseline: current load
- Data-driven inter-case feature encoding (DDICFE) [4] [5]

[4] Data-driven inter-case feature encoding, A.Senderovich, C.D.Francescomarino, and F.M.Maggi, “From knowledge driven to data-driven inter-case feature encoding in predictive process monitoring,” *Information Systems*, 2019.
[5] https://github.com/processmining-in-logistics/psm/tree/ppm/ppm/src/main/scala/org/processmining/scala/intercase

Experimental Results

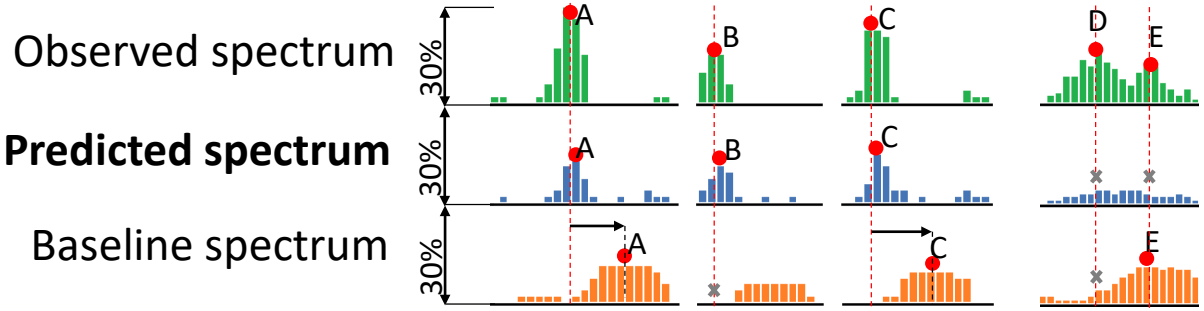
Problem Instance 1

Load on the X-Ray Screening Machines



Problem Instance 2

Extra re-circulation because of exits a1-a3 unavailability



Experiment	Approach	Model	RMSE, % of max
Sim. model	**PS**	**LR**	**12.1**
	DDICFE	**FFNN**	21.2
	Naïve baseline	-	35.8
Real system	**PS**	**LR**	**7.0**
	DDICFE	**FFNN**	7.6
	Naïve baseline	-	11.0

Experiment	Approach	Model	RMSE, % of max
Real system	**PS**	**FFNN**	**2.4**
	Naïve baseline	-	3.0

Conclusion and Future Work

- Processes: **cases influence each other**
- **Performance Spectrum**:
  - modelling a **variety of performance features over time**
  - capturing **inter-case dependencies**
- **Multi-Channel PS:** formulating a large class of **performance prediction problems** as a **regression problem**
- **Generic methodology** of solving this problem as a **ML task**

- Feasibility on **MHS** → apply to **Business Processes**
- Feature selection: **domain knowledge is required** → include a **formal process model**
- **Validation in practice** is missing → **higher accuracy** and **longer prediction horizon**

Thank you!