SAMSUNG

Multi-view Feature-based SSD Failure Prediction: What, When, and Why

Yuqi Zhang†, Wenwen Hao†, Ben Niu‡, Kangkang Liu‡, Shuyang Wang†, Na Liu†, Xing He†,
Yongwong Gwon∗, Chankyu Koh∗

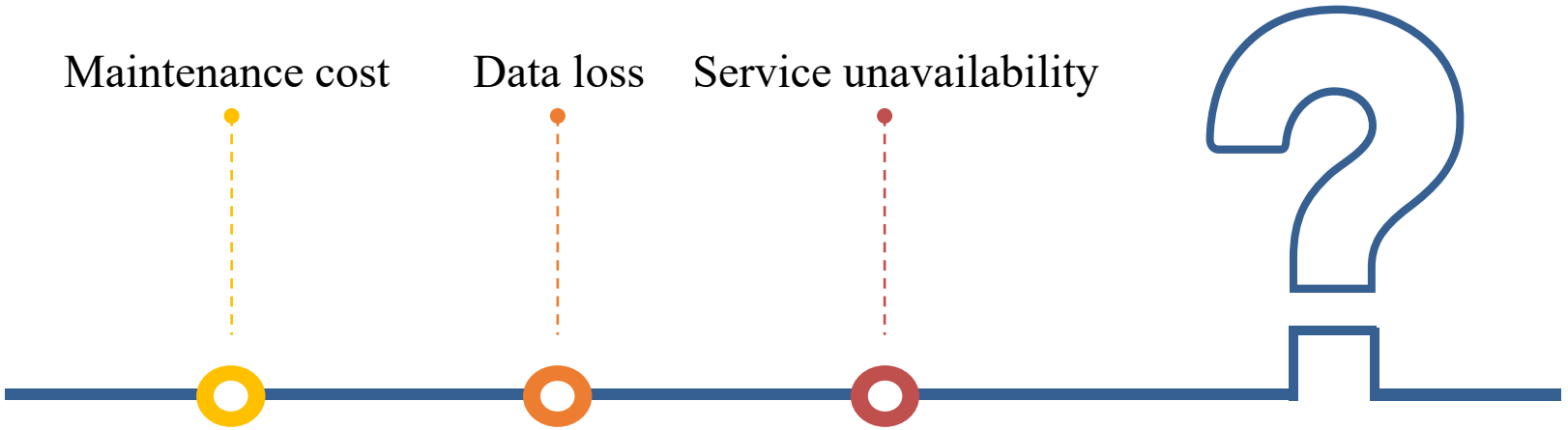
†Samsung R&D Institute China Xi'an, Samsung Electronics
‡Tencent
∗Samsung Electronics

2023 SRCX

Samsung R&D Institute China Xi'an(SRCX) established in 2013 and located in Xi'an High-tech Zone, is the only cutting-edge technology R&D institute of Samsung Electronics in western China.

Introduction

- **SSD failure**
  - Leads to a challenge with large-scale deployment of SSDs.
  - Brings an impact on storage performance and reliability in data centers.



Introduction

- **SSD failure prediction**
  - Recognized as a proactive fault tolerance mechanism
  - Failure/health classification with ML based on SSD monitoring logs
- **Remaining major challenges**
  - Failure pattern may be hidden in long-term monitoring data
  - Lack instructive suggestions for verifying and handling failures

long sequence
various length

internal failure?
urgent?
serious?

Introduction

- **Our work**
  - SSD failure analysis on monitoring data from Tencent data center

    - **70 million**+ Telemetry logs (SMART/customized attributes)
    - **300,000**+ Samsung SSDs
    - Failure and handing information are collected and findings are drawn

  - Multi-view and multi-task Random Forest (***"MVTRF"***) scheme is proposed based on findings:

    - **Multi-view** features with both short-term and long-term monitoring info to capture failures
    - **Multi-task** model to obtain detailed failure info for recommending handling measures

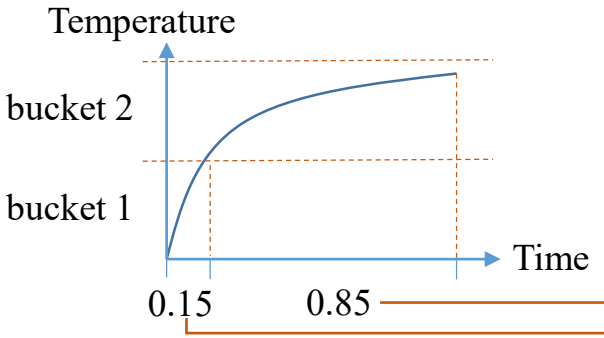
西安三星电子研究所 | SAMSUNG R&D INSTITUTE CHINA XIAN

Analysis and Design (1)

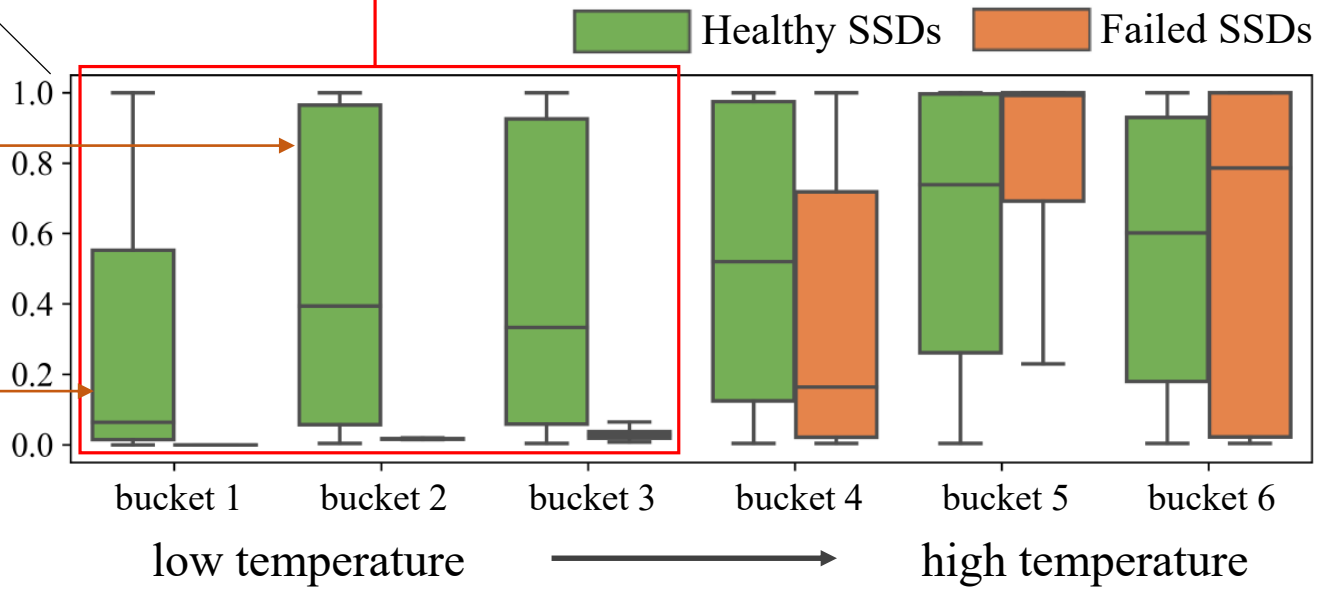
➢ **Finding:** The statistics of long-term Telemetry data of each SSD are useful for failure prediction

The proportions of long-term data of each SSD in each temperature bucket (range)

SSDs with shorter time at low temperature are more likely to fail



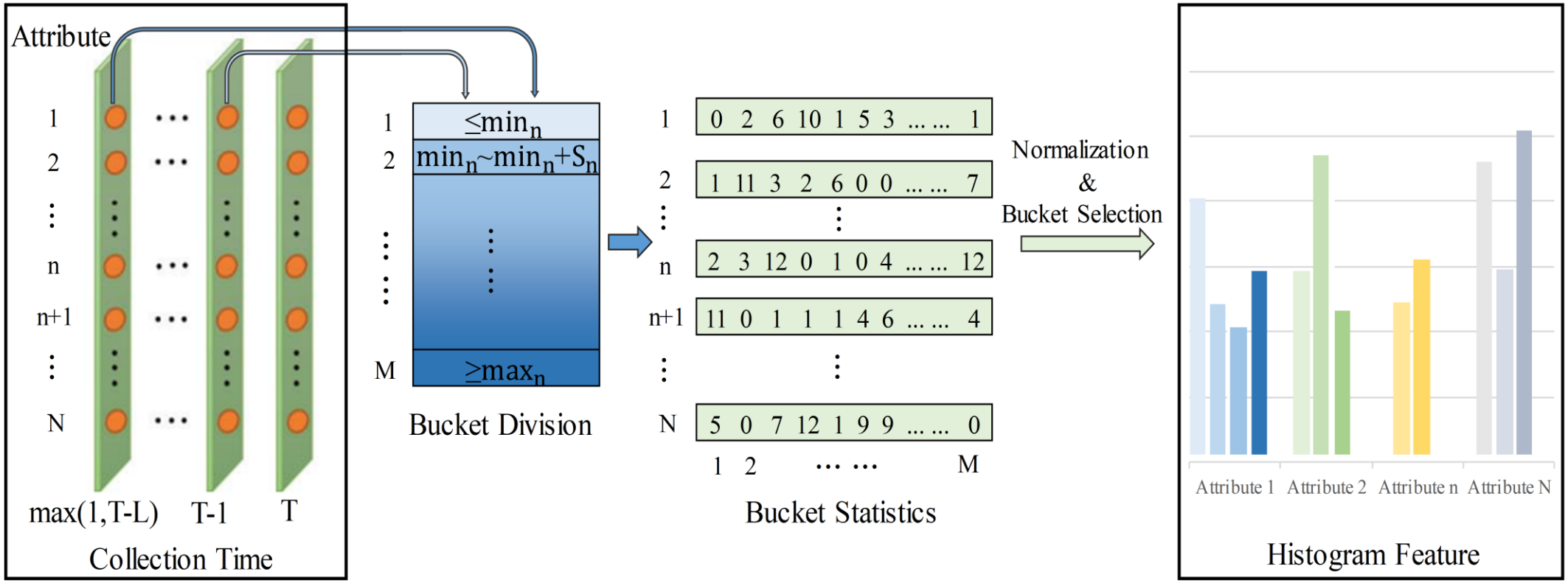
Example of proportions (Single SSD)



Analysis and Design (1)

➢ **Design: Histogram feature**

- Bucket statistics of long-term data for each attribute of each SSD
- Divide the long-term multiple values into buckets, count and normalize them



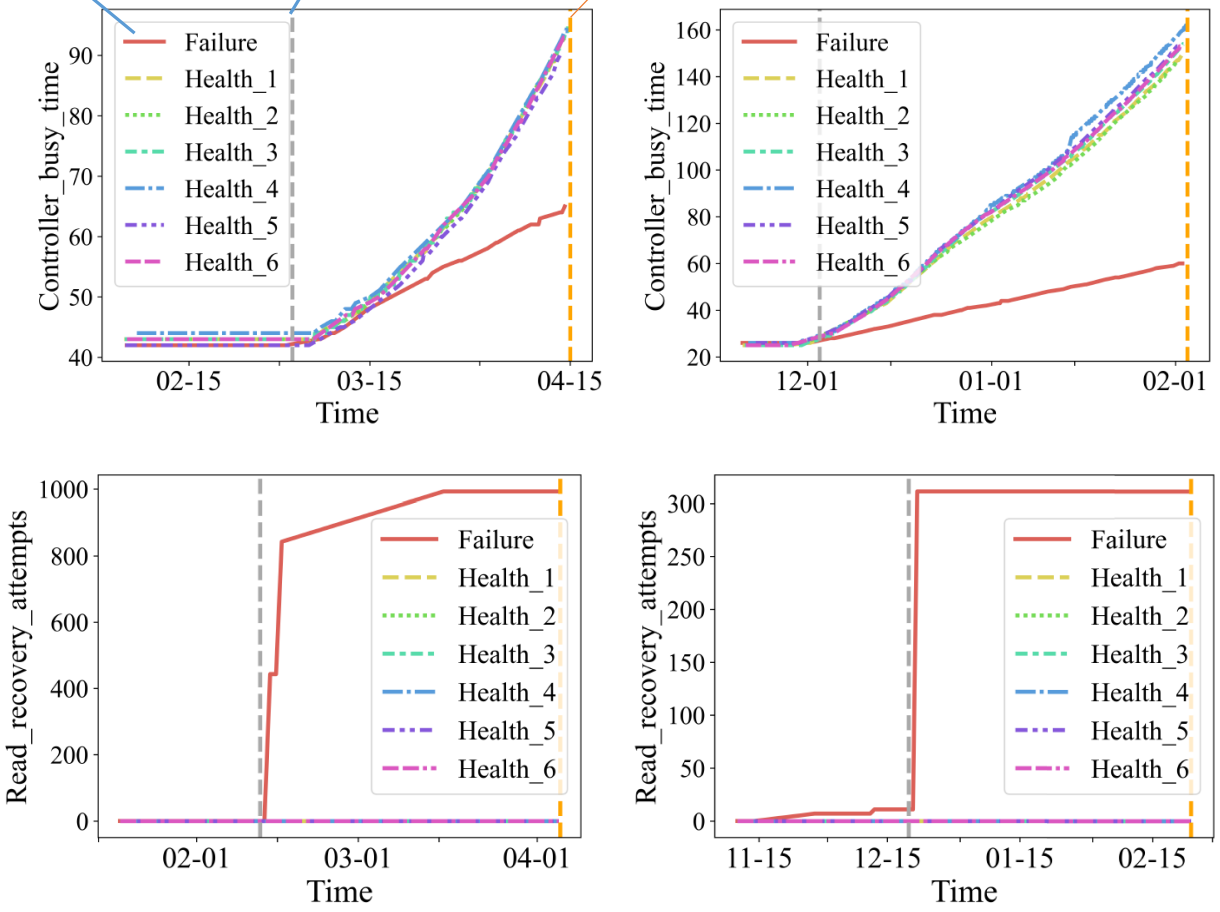
Analysis and Design (2)

➢ **Finding**: Some failure symptoms are hidden in long-term attribute trends

Failed SSD and healthy SSDs on the same server

Symptom-showing date

Failure-reporting date



Two Read Only failed SSDs show gradual growth rate of controller_busy_time before failure report

Two Health Check failed SSDs show rapid rise of read_recovery_attempts before failure report

Analysis and Design (2)

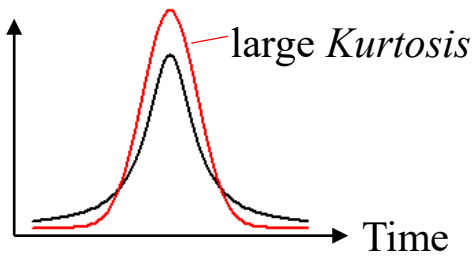
➢ **Design: Sequence-related feature**

- Characterize the fluctuation and trend for each attribute of each SSD

  1) *Coefficient of variation*: dispersion degree

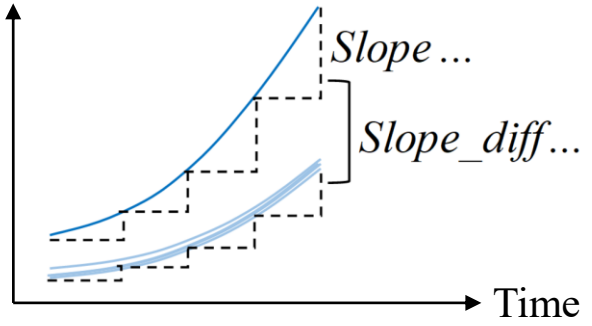
  2) *Kurtosis*: steepness of the trend

  3) *Slope*: changing trend



- Reflect the trend anomaly for each attribute of each SSD

  4) Differences in the above features with other SSDs on the same server



- Divide long-term data into segments and generate above features for each segment

Analysis and Design (3)

➢ **Finding**: Detailed failure information is significant for failure handling

- Collect failure descriptions and handling status from DC's failure list.
- Operators check failures manually (SSD? backplane? environment?)
- The urgency and measures can vary for different failure type/phenomenon

Failure type	Description	Urgency
Check Failed	Health or performance check failed	High
Cancelling I/O	NVMe cancelling I/O	Medium
PLP	Power loss protection test failed	Medium
SSD Drop	SSD cannot be detected by host	Medium
Fail Mode	Device fail mode	Medium
Media Error	Some data cannot be read correctly	Medium
Read Only	Unable to write data to SSD	Medium
Reliability Degradation	NVMe reliability degradation	Low

(8 failure types)

For this type, almost all failed SSDs were quickly replaced

For this type, the failed SSDs were handled slowly and a small number of them were replaced

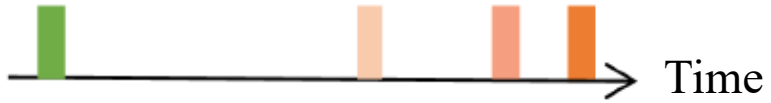
Analysis and Design (3)

➢ **Design: failure type and lifespan prediction and failure cause identification**

- Classification on 8 types of failures



- Classification on remaining lifespan



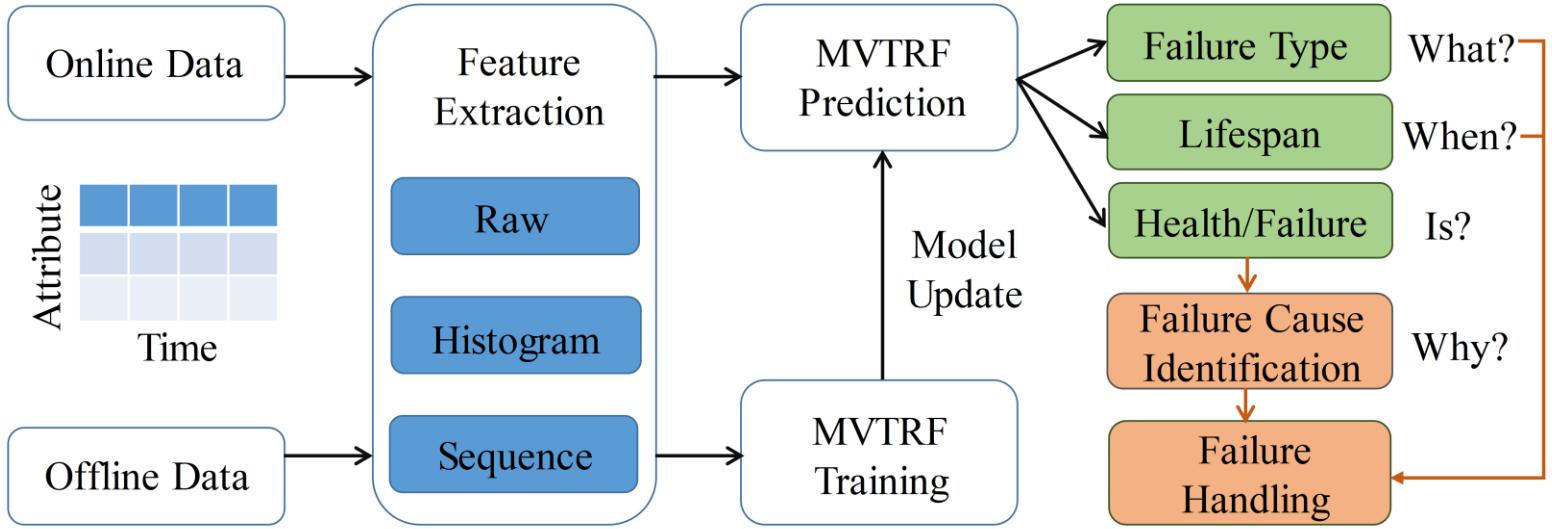
- Extract ML decision processes to reveal failure causes

media_errors > 1000
wear_leveling_count > 50
……

bad_tlp_count > 6000
……

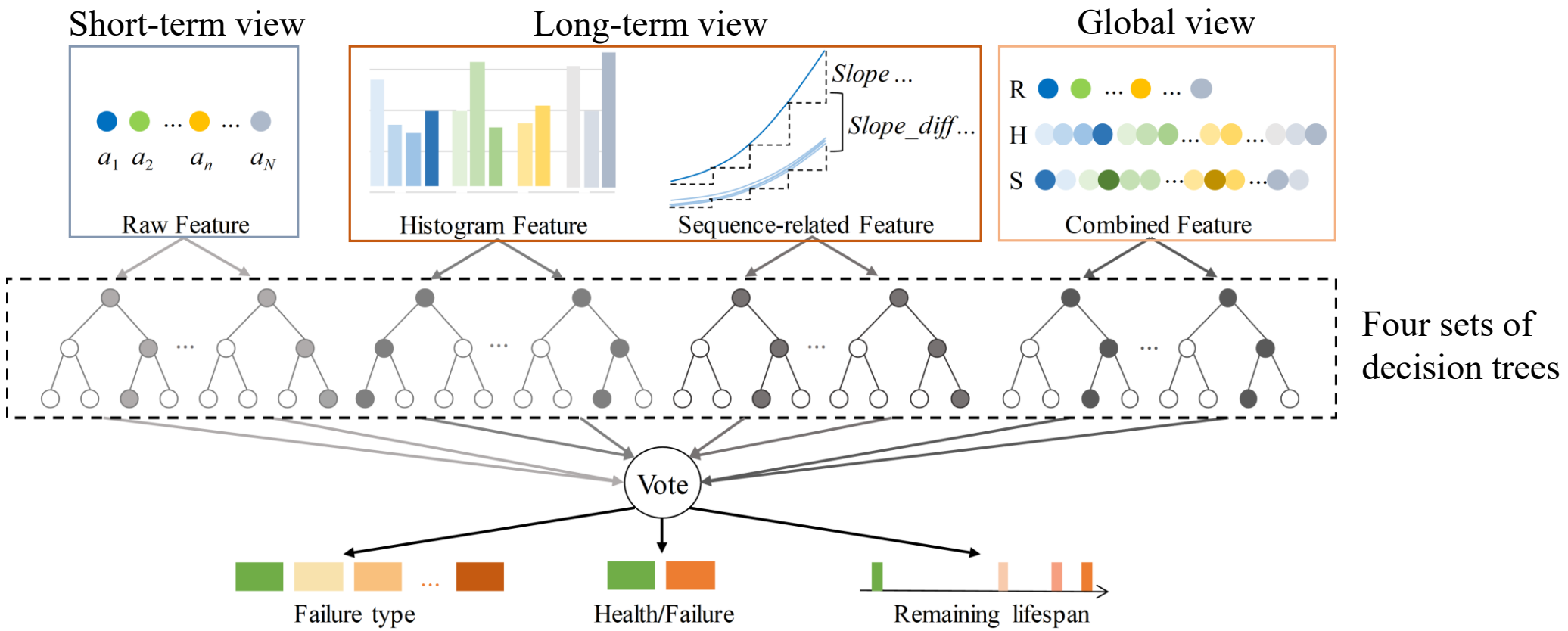
Implementation

- **Overall architecture includes …**
  - Multi-view feature extraction
  - Multi-view and multi-task Random Forest (MVTRF)
  - Health or failure/failure type/Lifespan/failure cause identification



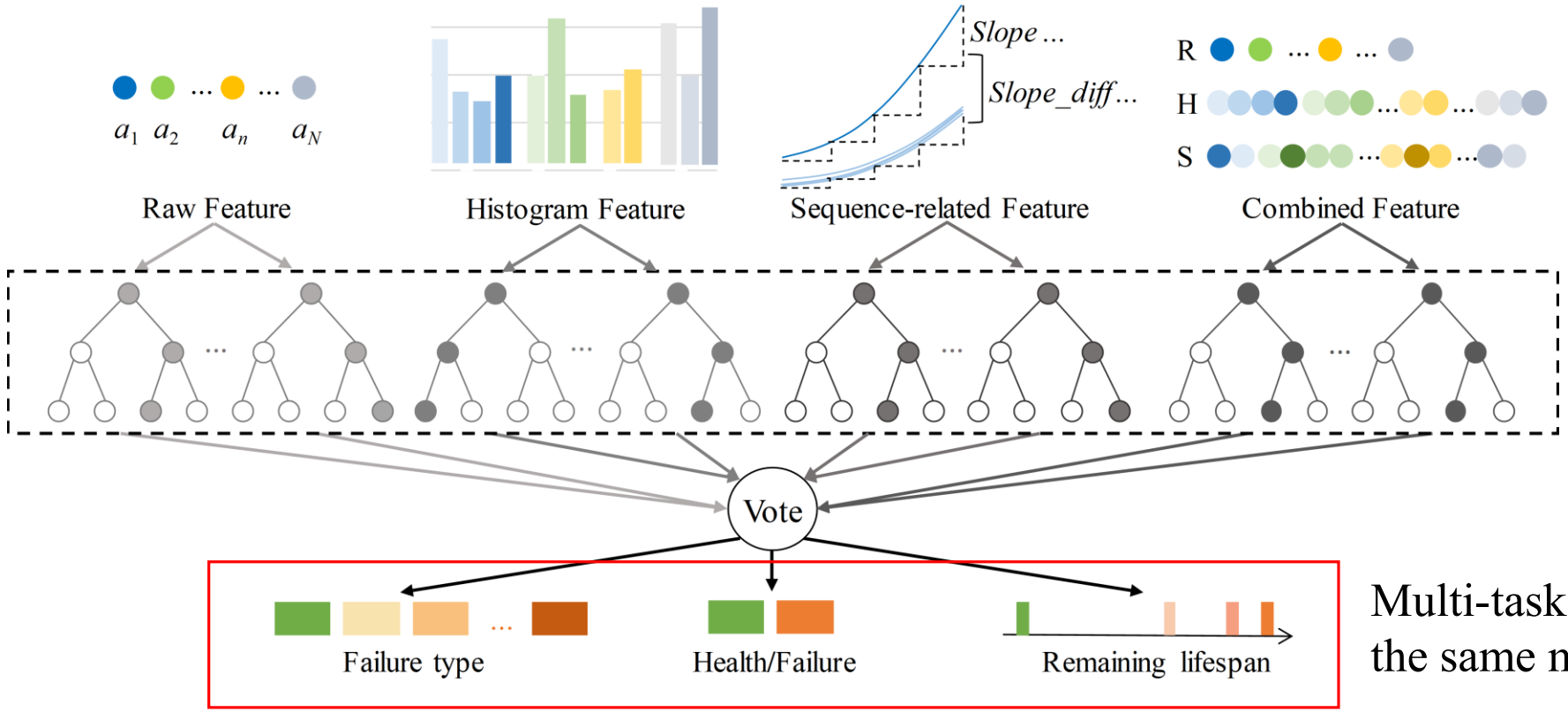
Implementation

- **MVTRF : Multi-view** and multi-task Random Forest
  - The four types of features are input into the four sets of decision trees respectively
  - Combining multiple views, all decision trees vote to get the final result



Implementation

- **MVTRF :** Multi-view and **multi-task** Random Forest
  - Besides Health/Failure, failure type and lifespan are predicted for failure handling
  - Accuracy would be improved with cost re by sharing model in related tasks



Multi-task with the same model

Implementation

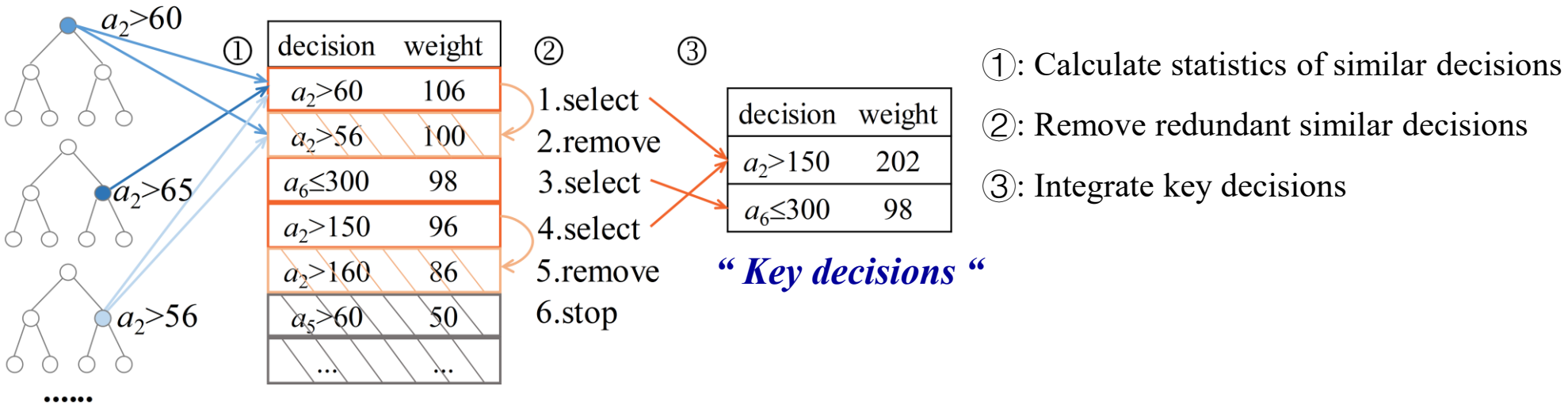
- **Failure cause identification**
  - Similar decision extraction (SDE) to reveal failure causes and verify the failure

- Similar decision's criteria :

1) The same feature and decision logic
2) Difference between decision thresholds is within 10%

- Extracted key decisions

Similar decisions that appear more frequently in multiple decision trees



Evaluation

- **Dataset setup**

Dataset	Experiment round	Train set (month)	Val set (month)	Test set (month)	
Samsung PM1733 (Tencent)	1	1–7th	8th	9th	Tencent Telemetry datasets
	2	1–6th	7th	8th	
Samsung PM9A3 (Tencent)	1	1–7th	8th	9th	
	2	1–6th	7th	8th	
MB1 (Alibaba)* (detailed model unkown)	1	1–22th	23th	24th	Alibaba public SMART dataset
	2	1–21th	22th	23th	
	3	1–20th	21th	22th	

- **Four metrics (P, R and F0.5 : Binary Classification)**
  - *Precision*: The proportion of true alarms (correctly predicted) to both true alarms and false alarms
  - *Recall*: The proportion of true alarms to all actual failed SSDs
  - *F0.5-Score*: $\frac{(0.5^2+1)\times Precision\times Recall}{0.5^2\times Precision+Recall}$ the harmonic average of precision and recall, where precision is weighted higher
  - *ROC_AUC*: A score reflecting the classification ability across discrimination thresholds

*Fan Xu, et al. General Feature Selection for Failure Prediction in Large-scale SSD Deployment, DSN 2021

西安三星电子研究所 | SAMSUNG R&D INSTITUTE CHINA XIAN

Evaluation

- **Comparison with existing schemes**
  - MVTRF shows better performance on all three datasets and significantly improves four metrics on average, which demonstrates its effectiveness and generalizability

Methods	PM1733 Tencent				PM9A3 Tencent				MB1 Alibaba				Average			
	P	R	F	AUC	P	R	F	AUC	P	R	F	AUC	P	R	F	AUC
Random Forest	0.58	0.31	0.48	0.69	0.75	0.33	0.60	0.75	0.56	0.37	0.51	0.87	0.63	0.34	0.53	0.77
Neural Network	0.63	0.14	0.36	0.58	**0.85**	0.31	0.58	0.61	0.72	0.46	0.64	**0.89**	0.73	0.30	0.53	0.69
Autoencoder	0.54	0.14	0.34	0.77	0.54	0.33	0.40	0.78	0.36	0.46	0.31	0.88	0.48	0.31	0.35	0.81
LSTM	0.36	**0.40**	0.36	0.69	0.52	0.25	0.28	0.62	0.63	0.61	0.62	0.87	0.50	0.42	0.42	0.73
MVTRF(Ours)	**0.90**	**0.40**	**0.72**	**0.81**	0.70	**0.42**	**0.61**	**0.83**	**0.89**	**0.76**	**0.86**	0.86	**0.83**	**0.53**	**0.73**	**0.83**

(P: precision; R: recall; F: F0.5-Score; AUC: ROC_AUC)

Evaluation

- **Discussion on multi-view features**
  - Raw, histogram, and sequence features with different views perform better on different metrics

Method	Precision	Recall	F0.5-Score	ROC_AUC
RF + Raw	0.61	0.34	0.52	0.69
RF + Histogram	**1.00**	0.17	0.48	0.72
RF + Sequence	0.50	0.25	0.38	**0.81**
RF + Combined	0.83	0.37	0.66	0.78
NN + Combined	0.79	0.17	0.39	0.74
AE + Combined	0.88	0.14	0.40	0.77
MVTRF	0.90	**0.40**	**0.72**	**0.81**

(RF: Random Forest; NN: Neural Network; AE: Autoencoder)

Evaluation

- **Discussion on multi-view features**

  - Combined features contain much information and this may lead to overfitting in model training
  - MVTRF combining multiple views reduces the overfitting caused by only using combined features

Method	Precision	Recall	F0.5-Score	ROC_AUC
RF + Raw	0.61	0.34	0.52	0.69
RF + Histogram	**1.00**	0.17	0.48	0.72
RF + Sequence	0.50	0.25	0.38	**0.81**
RF + Combined	0.83	0.37	0.66	0.78
NN + Combined	0.79	0.17	0.39	0.74
AE + Combined	0.88	0.14	0.40	0.77
MVTRF	0.90	**0.40**	**0.72**	**0.81**

(RF: Random Forest; NN: Neural Network; AE: Autoencoder)

Evaluation

- **Multi-task learning and prediction**
  - Multi-task with model sharing boosts the failure prediction performance
  - The failure type and lifespan prediction with multi-task model achieves the highest accuracy rate (0.95 and 0.55 respectively)

Method		P	R	F	AUC	Type Acc	Lifespan Acc
RF + Raw	Single-task	0.58	0.31	0.48	0.69	0.88	0.44
	Multi-task	0.61	0.34	0.52	0.69	0.93	0.53
MVTRF	Single-task	0.83	0.37	0.66	0.79	0.93	0.47
	Multi-task	**0.90**	**0.40**	**0.72**	**0.81**	**0.95**	**0.55**

(P: precision; R: recall; F: F0.5-Score; AUC: ROC_AUC; Acc: accuracy rate)

Evaluation

- **Similar decision extraction (SDE)**
  - False alarms are reduced with the combination of key decisions which can represent the major decision process

Key decisions of an SSD failure

	Key decision	Feature type	Weight	False alarms
①	$media_errors_slope > 126.44$	Sequence	122	3 (①)
②	$media_errors_bkt0 <= 0.99$	Histogram	120	3 (① - ②)
③	$temperature_kurt <= -1.11$	Sequence	117	1 (① - ③)
④	$media_errors > 6015.5$	Raw	113	1 (① - ④)
⑤	$wear_leveling_max_kurt > -0.047$	Sequence	96	0 (① - ⑤)

- **Key decisions perform as well as all decisions in predicting failed SSDs**

	Decision NO.	Precision	Recall
All decisions	53663	0.90	0.40
Key decisions	49	0.90	0.40



x1000 times

Conclusion

- Combining multi-view features (involving both short-term and long-term information) can find more failures and reduce false alarms
- Multi-task model is effective in predicting health/failure, failure type, and lifespan simultaneously, which provides more suggestions for operators to deal with failures
- Extracting the decision process of machine learning helps to reveal failure causes and thus verify the failure

SAMSUNG

2023 SRCX

THANK YOU

Samsung R&D Institute China Xi'an(SRCX) established in 2013 and located in Xi'an High-tech Zone, is the only cutting-edge technology R&D institute of Samsung Electronics in western China.

西安三星电子研究所 | SAMSUNG R&D INSTITUTE CHINA XIAN