

Log-based Predictive Maintenance

Ruben Sipos, Cornell U.

Dmitriy Fradkin, Siemens

Fabian Moerchen, Amazon

Zhuang (John) Wang, Skytree

(work performed while all authors were employed by Siemens Research in Princeton, NJ)

Log-based Predictive Maintenance

- Helps determine the running condition of in-service equipment to predict when and where repairs should be performed

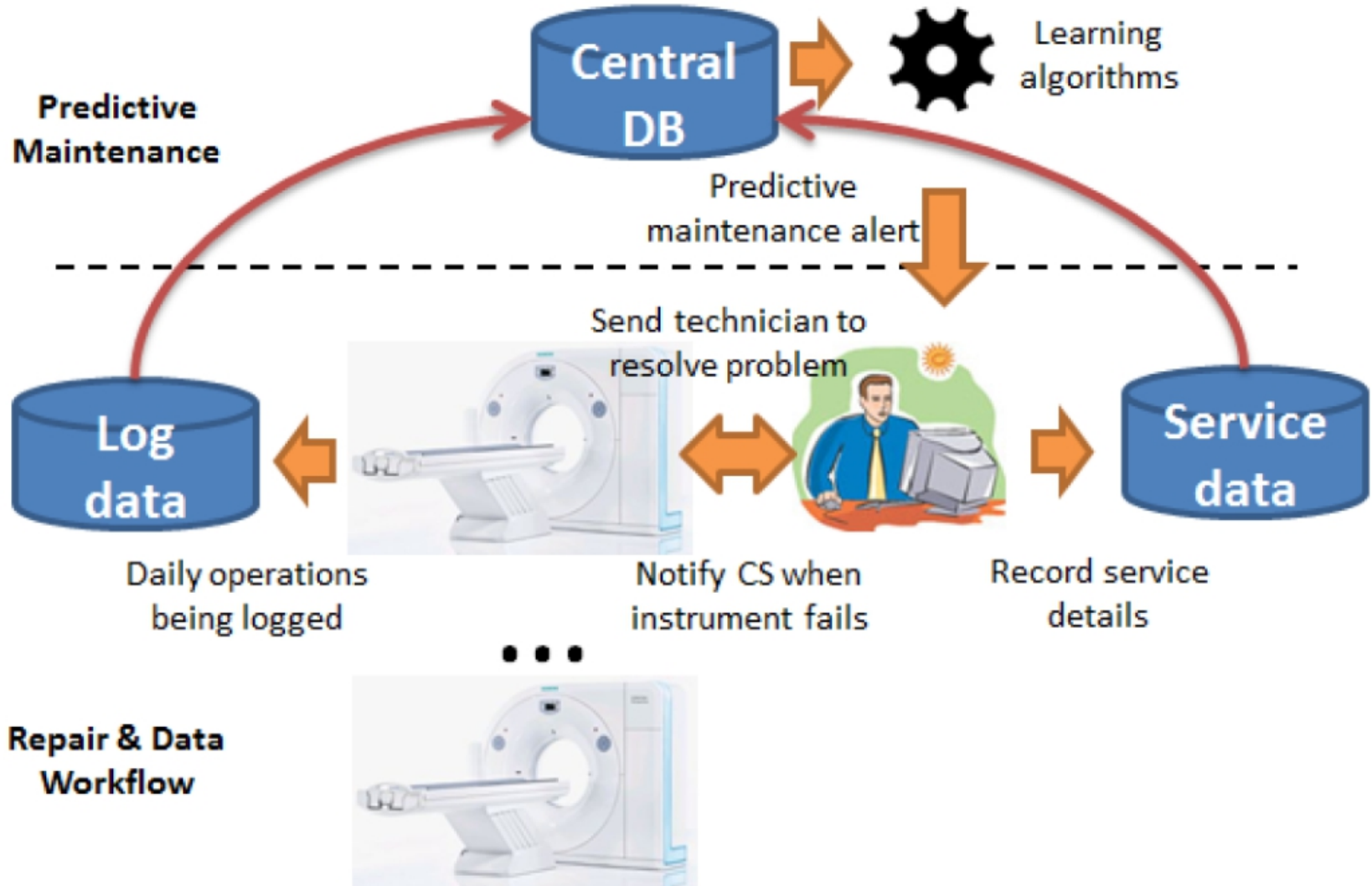


Logs from various software applications record detailed operations of equipment

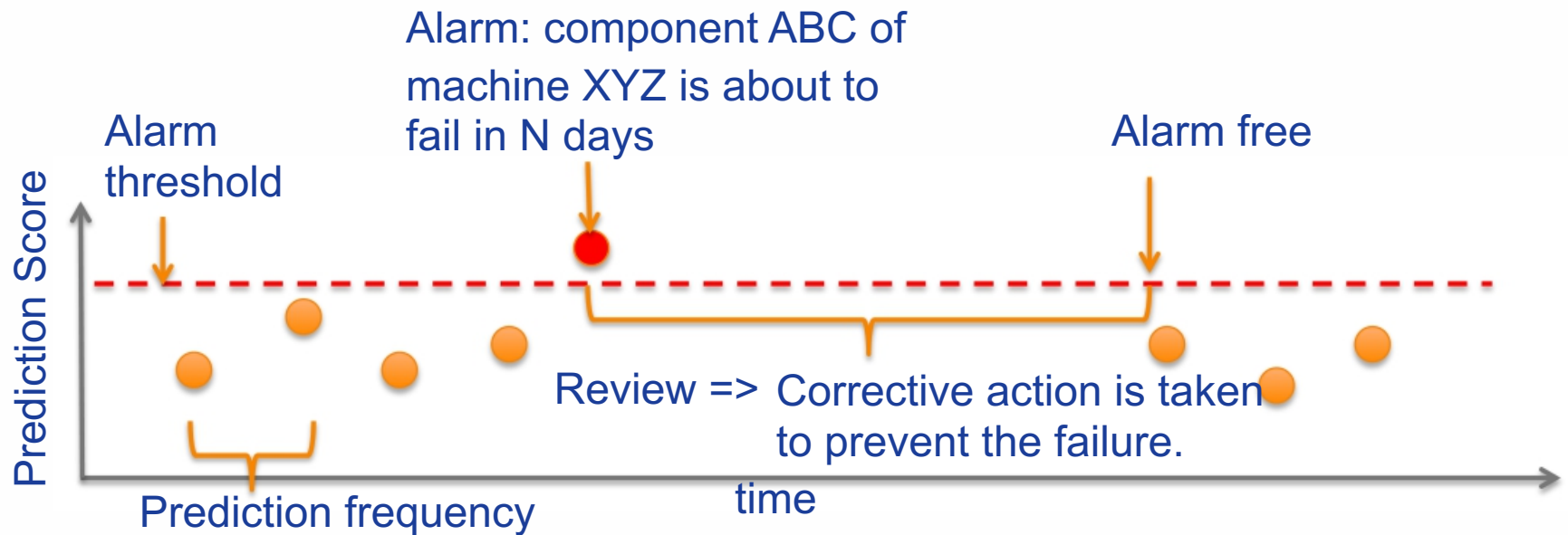
```
2012-09-08 21:21:09 CT_ITH 2 IRS scan information: Topo tot Rdg: 6460, defects: HV-Drop 0, H1
2012-09-08 21:21:07 CT_IS 320 First data from IRS arrived in IS
2012-09-08 21:21:06 CT_CHR 107 SKIP button is hidden (bolus exam not started:active entry is not n
2012-09-08 21:21:06 CT_CHR 107 CANCEL button is dimmed (scanstate == UiScanning)
2012-09-08 21:21:06 CT_CHR 107 CANCELMOVE button is hidden (scanstate == UiScanning)
2012-09-08 21:21:06 CT_STC 93 Zero button has been locked.
2012-09-08 21:21:06 CT_ISV 49 Timer was started waiting for ..... some answer from IRS Recc
2012-09-08 21:21:06 CT_CHR 107 Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT REGISTRATION EXAM#
2012-09-08 21:21:06 CT_ISV 49 Timer was started waiting for ..... SCAN_DONE message for 77E
2012-09-08 21:21:05 CT_ACU 3037 Control info ACU (E c0 03 25 d4 4a 00 00)
2012-09-08 21:21:03 CT_ACU 3037 Control info ACU (E c0 03 25 d3 7e 00 00)
2012-09-08 21:21:02 CT_TG 9 A new planning image of type CLEAR SEGMENT with series loid EMPTY
2012-09-08 21:21:02 CT_MSM 211 (+) Receiving MeasStart request (scan 0 of range 0)
2012-09-08 21:21:02 CT_SSQ 493 received IS-Notification: SsqMeasISClient::onIrsStarted / IS_MsgId
2012-09-08 21:21:02 CT_SIS 120 sStart button released
2012-09-08 21:21:02 CT_SIS 85 Multiplexer: the START-button on control box has been released.
2012-09-08 21:21:02 CT_ISV 46 Receiving IsvMsgStartAck message of size 80 Bytes from IRS.
2012-09-08 21:21:02 CT_CHR 107 SKIP button is hidden (scanstate == UiStartedTopobolus exam not st
2012-09-08 21:21:02 CT_CHR 107 CANCEL button is dimmed (scanstate == UiStartedTopo)
2012-09-08 21:21:02 CT_CHR 107 CANCELMOVE button is hidden (scanstate == UiStartedTopo)
2012-09-08 21:21:02 CT_CHR 107 LOAD button is hidden (suspend-button is shown)
2012-09-08 21:21:02 CT_CHR 107 SUSPEND button is ENABLED
2012-09-08 21:21:02 CT_CHR 107 Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT REGISTRATION EXAM#
2012-09-08 21:21:02 CT_ISV 45 Sending IsvMsgStart message of size 72 Bytes to IRS.
```

Mining logs helps in detecting potential issues in advance.

Maintenance Service Process and Data Flow

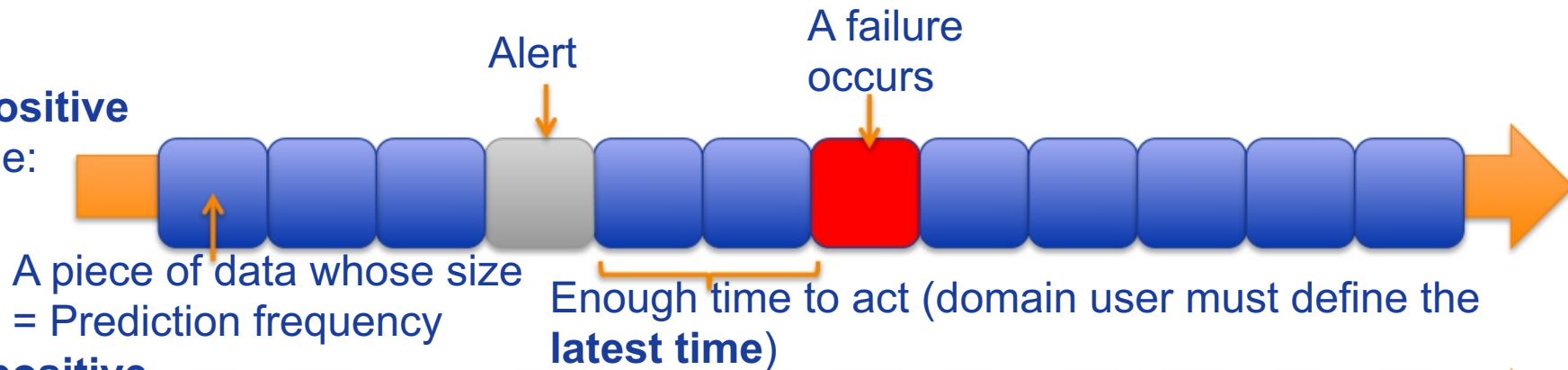


How is Model Used in Real-life?

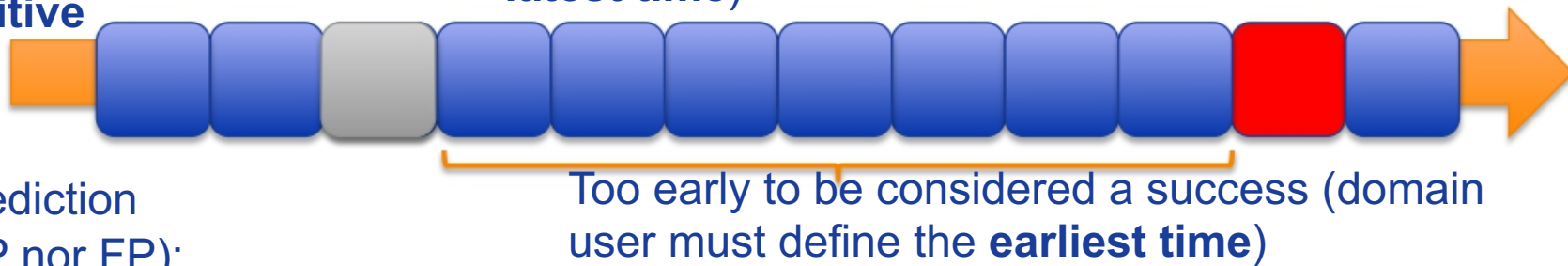


How is Model Evaluated in Real-life?

True positive
example:



False positive
example:



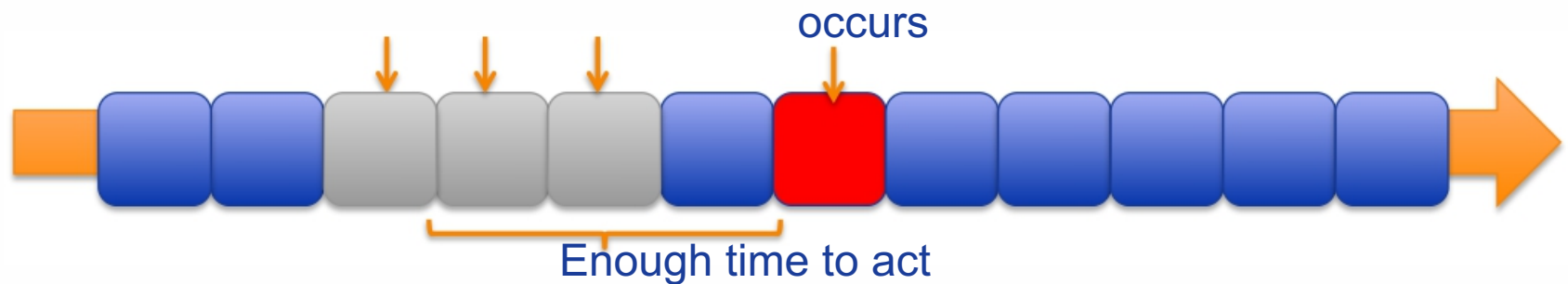
Invalid prediction
(neither TP nor FP):



How is Model Evaluated in Real-life? (cont.)

Counting of TP/FP:

3 true alerts but only count 1 TP (1



2 false alerts but only count 1 FP (1 Responsive Duration can only have at most 1 FP)



- **Recall:** TP/ All failures
- **Precision:** TP/ (TP+ FP)
- **Predictive-Maintenance-based AUC (PM-AUC):** computed like regular PR-AUC but using the above defined TP and FP.

Application Event Logs: the X

- Logs contain every detail of how an equipment was being operated.
- **Logs contain complicated data types.**

	Time stamp	Eventcode	Message text
event	2012-09-08 21:21:09	CT_ITH 2	IRS scan information: Topo tot Rdg: 6460, defects: HV-Drop 0, H1
	2012-09-08 21:21:07	CT_IS 320	First data from IRS arrived in IS
Temporal sequence	2012-09-08 21:21:06	CT_CHR 107	SKIP button is hidden (bolus exam not startedactive entry is not n
	2012-09-08 21:21:06	CT_CHR 107	CANCEL button is dimmed (scanstate == UiScanning)
	2012-09-08 21:21:06	CT_CHR 107	CANCELMOVE button is hidden (scanstate == UiScanning)
	2012-09-08 21:21:06	CT_STC 93	Zero button has been locked.
	2012-09-08 21:21:06	CT_ISV 49	Timer was started waiting for some ar IRS Recc
	2012-09-08 21:21:06	CT_CHR 107	Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT EXAM#
	2012-09-08 21:21:06	CT_ISV 49	Timer was started waiting for SCAN_D for 776
	2012-09-08 21:21:06	CT_ACU 3037	Control info ACU (E c0 03 25 d4 4a 00 00)
	2012-09-08 21:21:06	CT_ACU 3037	Control info ACU (E c0 03 25 d3 7e 00 00)
	2012-09-08 21:21:02	CT_TG 9	A new planning image of type CLEAR SEGMENT with series loid EMPTY
	2012-09-08 21:21:02	CT_MSM 211 (+)	Receiving MeasStart request (scan 0 of range 0)
	2012-09-08 21:21:02	CT_SSQ 493	received IS-Notification: SsqMeasISClient::onIrsStarted / IS_MsgIc
	2012-09-08 21:21:02	CT_SIS 120	sStart button releaseds
	2012-09-08 21:21:02	CT_SIS 85	Multiplexer: the START-button on control box has been released.
	2012-09-08 21:21:02	CT_ISV 46	Receiving IsvMsgStartAck message of size 80 Bytes from IRS.
Categorical variables	2012-09-08 21:21:02	CT_CHR 107	Button(s) is hidden (scanstate == UiStartedTopobolus exam not st
	2012-09-08 21:21:02	CT_CHR 107	Button(s) is dimmed (scanstate == UiStartedTopo)
Unstructured text	2012-09-08 21:21:02	CT_CHR 107	Button(s) is hidden (scanstate == UiStartedTopo)
	2012-09-08 21:21:02	CT_CHR 107	LOA button is hidden (suspend-button is shown)
	2012-09-08 21:21:02	CT_CHR 107	SUSP button is ENABLED
	2012-09-08 21:21:02	CT_CHR 107	Button(s) disabled:#CLOSE PATIENT#EXAM#PATIENT REGISTRATION EXAM#
	2012-09-08 21:21:02	CT_ISV 45	Sending IsvMsgStart message of size 72 Bytes to IRS.
	2012-09-08 21:21:02	CT_ISV 45	Sending IsvMsgStart message of size 72 Bytes to IRS.
	2012-09-08 21:21:02	CT_ISV 45	Sending IsvMsgStart message of size 72 Bytes to IRS.
	2012-09-08 21:21:02	CT_ISV 45	Sending IsvMsgStart message of size 72 Bytes to IRS.
	2012-09-08 21:21:02	CT_ISV 45	Sending IsvMsgStart message of size 72 Bytes to IRS.
	2012-09-08 21:21:02	CT_ISV 45	Sending IsvMsgStart message of size 72 Bytes to IRS.

Encoded
Time
series

Service Data: the Y

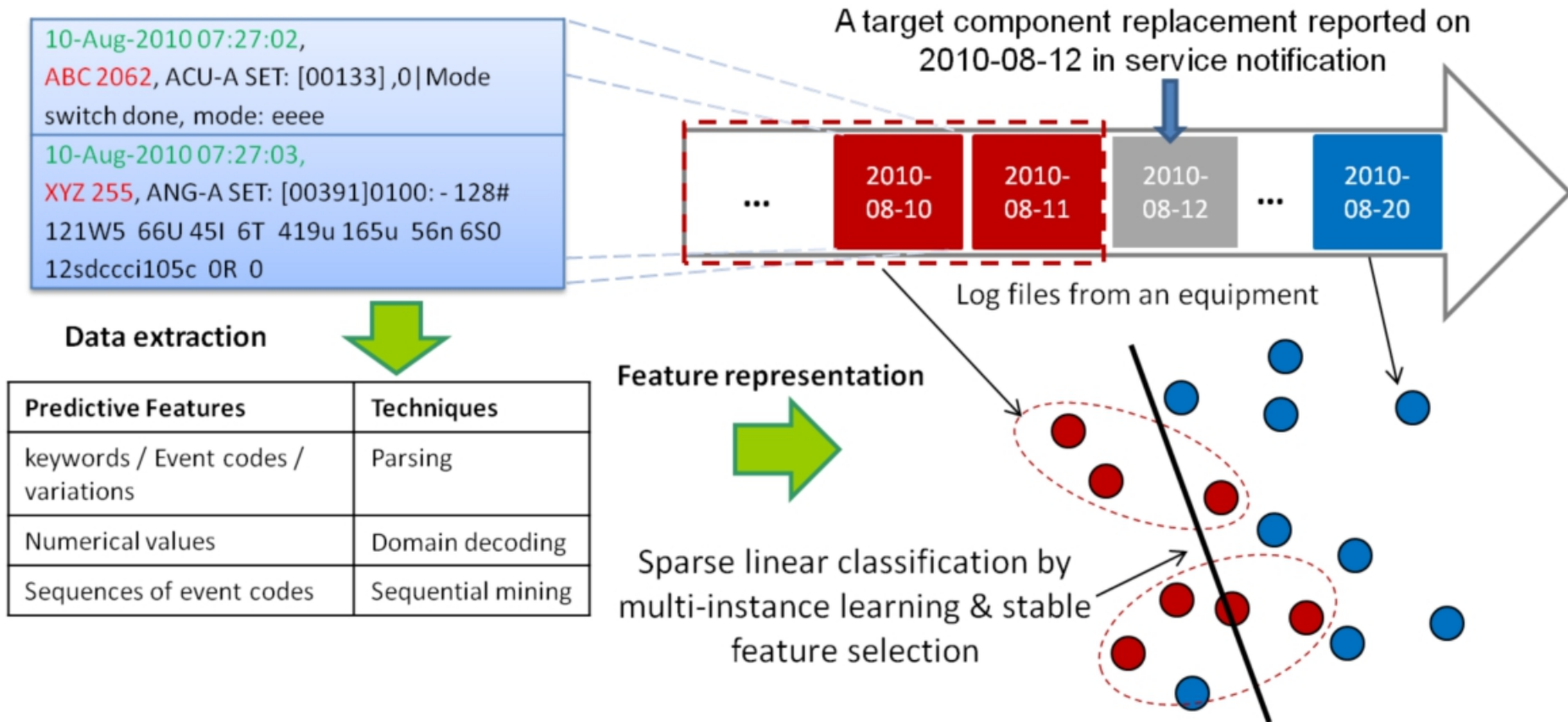
- Service data records details of performed services:
 - service open/close date
 - equipment id
 - component replacement
 - ...

Equipment	Notification Number	Notification date	Notification time	Text short	Ausfalldatum ab	Ausfalzeit ab	Ausfalldatum bis	Ausfalzeit bis	Division	IPB	Material		BCKF I	BCKF	MVKZ	
	Gesamtergebnis												5,283	5,071,584.87	3,301	
14003	728100458470	07.09.2007	12:41:33	GENERATOR B	07.09.2007	12:41:33	07.09.2007	23:00:00	01	AXA	05784506	MEGALIX Cat 125	1	17,778.75		
26	728100543184	21.02.2008	15:38:47	SOFTWARE A	21.02.2008	15:38:47	06.03.2008	22:38:34	01	AXA	04776063	HR Helflex Optik d	1	2,698.06		
27											05896852	CCR Board D71	1	1,547.29		
28											07149979	Kamera Head TH	1	5,081.49		
29	728100546276	27.02.2008	11:26:07	SOFTWARE H	27.02.2008	11:26:07	22.03.2008	02:56:31	01	AXA	07128866	CPU_D10_IS_OPE	1	1,077.50		
30	728100963980	15.12.2009	11:22:05	X-RAY NO X-R	15.12.2009	11:22:05	15.12.2009	20:00:00	01	AXA	05997617	PCA QUAD DSP1	1	1,077.50		
31	14004	400101791908	19.01.2007	16:38:12	CSE/27068 RE	19.01.2007	16:38:12	25.01.2007	20:24:00	01	AXA	01192124	ABSTANDSHALTE	2	5.52	
32	400101828175	28.02.2007	11:50:16	NO XRAY,ROD	28.02.2007	11:50:16	29.08.2007	16:30:00	01	AXA	04775990	D90 Ein/Aus (Log	1	534.44		
33											05246264	CBL Signl Comm D	4	206.88		
34											05998377	ASIM Host 3 Modul	1	2,432.56		
35											07325579	ASIM MPAC Mast	1	2,870.46		
36											07325900	ASIM PS2 WITH M	1	4,999.60		
37											07326080	PCA Power ON 2	1	724.08		
38	400101917799	05.06.2007	10:00:09	B-PLANE CAR	05.06.2007	10:00:10	09.08.2007	13:10:00	01	AXA	08485590	Potentiometer 354	2	108.62		
39											10140940	Schaltleiste Bx0B	2	698.22		
40	400101950864	11.07.2007	17:56:41	CD WRITER DC	11.07.2007	17:56:41	26.07.2007	16:40:00	01	AXA	01768535	WANDLER,AC-DC	1	208.17		
41											10051975	Adapter IDE to US	2	56.90		
42											10051978	Drive DVD/CD RW	1	140.94		
43	400102027866	03.10.2007	12:24:11	A-PLANE FLUX	03.10.2007	12:24:11	04.10.2007	09:40:00	01	AXA	07721603	FS Mono Artis Tie	1	1,383.51		
44	400102247951	05.03.2008	17:08:54	ECC RAIL CLA	05.03.2008	17:08:54	06.03.2008	11:30:00	01	AXA	04775735	Zubehoer-Schient	1	76.72		
45	400102260600	13.03.2008	10:33:50	TUBE STARTU	13.03.2008	10:33:51	14.03.2008	10:40:00	01	AXA	07124139	Anlassgeraet N75	1	2,736.85		
46	400102598104	21.10.2008	15:57:59	CSE COLE REF	21.10.2008	15:57:59	28.11.2008	09:10:00	01	AXA	04776063	HR Helflex Optik d	1	2,775.64		

- Not always correct but reasonable assumption:
 - a component replacement is the consequence of the component failure

We can use historical service data to correlate any known component failure with its corresponding equipment, time and relevant logs.

The Methodology



We propose a simple but effective algorithm for the resulting MIL problem with imbalanced labels and high-dimensional features

Build the MIL dataset:

1. Parse daily event logs to extract features.
2. Parse service notifications to extract known failure information.
3. Create bags for MIL learning.
 - (a) Group daily instances into bags using desired interval lengths. (e.g. one week)
 - (b) Label bags with known failures as positive and the rest as negative.

Train the model:

1. In the training phase, transform the MIL dataset.
 - (a) Each instance in a negative bag is transformed into a negative example.
 - (b) For each positive bag we create a positive meta example using the average of all bag's instances.

2. Feature selection.

- (a) Create multiple subsets by randomly subsampling transformed negative bags and including all positive meta examples.
- (b) Learn a sparse linear classifier on each subset.
- (c) Average weights from all runs and select features with the highest absolute weights.

- It is preferable to have an interpretable model so that experts are able to review it. This allows for incorporation of expert feedback into the modeling process.
- We use methods with L1 regularization to build sparse linear model which the domain experts can easily review.
 - Such a model consists of a weighted sum of a relatively small numbers of predictive features, with the weights specifying their precise contributions to the decision.
 - Furthermore, known noisy or irrelevant features specified by experts can be easily excluded from the new modeling process.

3. Train the final model.

- (a) Use the subset of features obtained in the previous step.
- (b) Learn the final model by using all the data.

Joining all the newly created examples from all the bags into a single training dataset $D' = \{(x_j, y_j), j = 1, \dots, M\}$, we formulate the learning problem as a L1-regularized SVM optimization, $\min_w \frac{\lambda}{2} \|w\|_1^2 + \sum_j \max\{1 - y_j w^T x_j, 0\}$, where $\lambda > 0$ is a user-specified regularization parameter. The optimization can be efficiently solved by Liblinear [9].

Test/Apply the model:

1. In evaluation, predict a bag as positive if the prediction score of a bag instance hits a given threshold.
2. In deployment, trigger an alert if the prediction score of a new daily log hits the pre-defined threshold.

Domain-based Evaluation Results

- Data

	Dataset C	Dataset A
# instances	181,945	422,560
# failures	88	108
# features	6,664	14,367

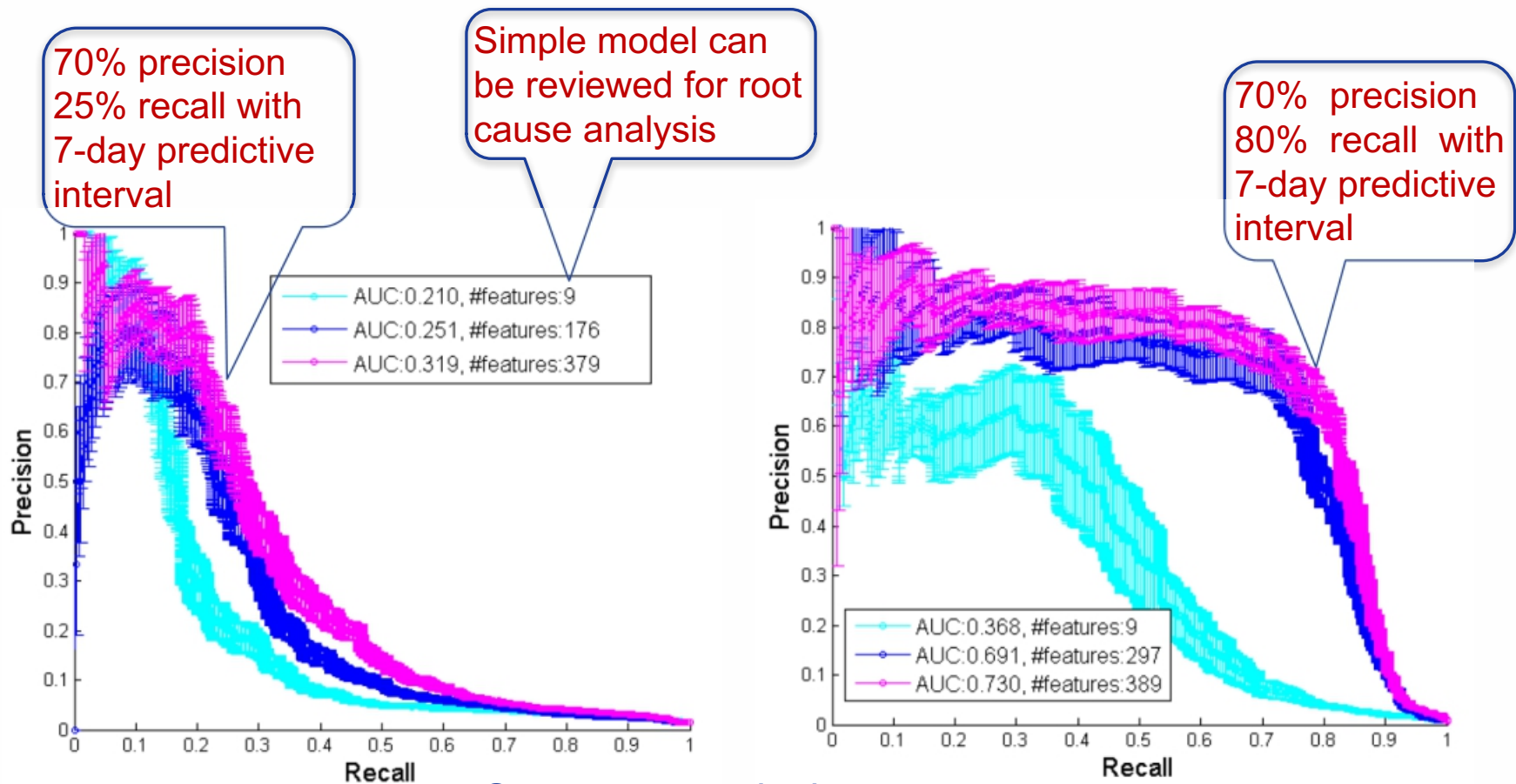
- Competing MIL algorithms

- AllInstance (ICML, 2005)
- Aggregated
- MILES (PAMI, 2006)
- MI-SVM (NIPS, 2003)

PM-AUC comparison (bag-level 5 CV with stratified sampling):

	<i>C</i>	<i>A</i>
random	0.037 (0.004)	0.017 (0.003)
AllInst.	0.293 (0.014)	0.620 (0.013)
Agg.	0.174 (0.013)	0.498 (0.016)
MILES	0.170 (0.011)	0.427 (0.117)
MI-SVM	0.216 (0.038)	0.700 (0.014)
Ours	0.319 (0.015)	0.730 (0.011)

PM-ROC with Different Model Complexity



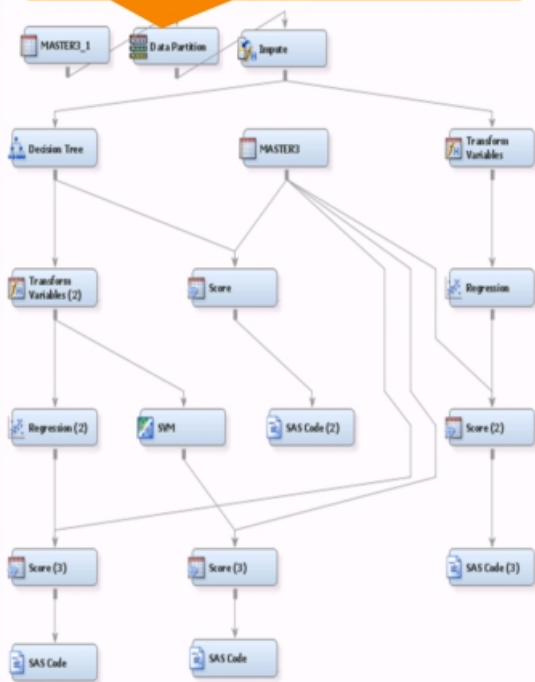
See more results in paper:

- On Utility of Single Features
- Local vs Global Models

Deployment

§ ■ Guided the client in implementing the approach on SAS platform.

Re-implemented the approach in SAS EM workflow



Portal Page
Decision Support Report

Bookmark Publish E-mail

Show only required items (denoted by *)

General

*Machine
54012

*Period
60

*Start Date
Current day of last year (October 19, 2011)

*Model
CT_P30_Tube_VE10C

User portal for decision support



alert scores

Predicted alerts for all the monitored equipments stored in a central DB

Severity	Dup.	Time Received	Machine	Model	alertday	DSReport	Message
Normal	2	05:54:22 05.09.12	CT8377520-55092	CT_P30_Tube_VD10A	2012.08.26	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 83%, FA: 17%, I
Normal	5	05:54:22 05.09.12	CT8377520-55092	CT_P30_Tube_VC11A	2012.08.24	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 50%, FA: 50%,
Normal	45	16:43:27 27.08.12	CT8377520-55140	CT_P30_Tube_VC11A	2012.08.23	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 50%, FA: 50%,
Normal	4	08:52:25 04.09.12	CT8377520-54876	CT_P30_Tube_VC11A	2012.08.15	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 50%, FA: 50%,
Normal	2	05:54:00 05.09.12	CT8377520-54586	CT_P30_Tube_VC11B	2012.08.14	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 80%, FA: 20%,
Normal	2	05:54:00 05.09.12	CT8377520-54586	CT_P30_Tube_VC11A	2012.08.14	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 50%, FA: 50%,
Normal	2	08:52:08 04.09.12	CT8377520-54432	CT_P30_Tube_VC11A	2012.07.27	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 50%, FA: 50%,
Normal	47	16:43:01 27.08.12	CT8377520-55526	CT_P30_Tube_VB10B	2012.07.24	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 53.9%, FA: 48
Normal	48	21:52:21 27.08.12	CT8377520-55324	CT_P30_Tube_VD10A	2012.07.21	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 83%, FA: 17%, I
Normal	97	21:52:21 27.08.12	CT8377520-54951	CT_P30_Tube_VD10A	2012.07.21	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 83%, FA: 17%, I
Normal	195	21:52:09 27.08.12	CT8377520-54734	CT_P30_Tube_VD10A	2012.07.12	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 83%, FA: 17%, I
Normal	42	21:51:55 27.08.12	CT8377520-54140	CT_P30_Tube_VB10A	2012.07.10	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 63.5%, FA: 36.!
Normal	1	05:54:15 05.09.12	CT8377520-54813	CT_P30_Tube_VC11B	2012.07.09	https://usmiv1bas05.www05...	Tube Failure Predicted: Precision: 80%, FA: 20%,

Thank you for your time! Questions?

End-to-end Predictive Maintenance App available at Skytree, Inc