ML-Assistant For Human Operators in Processing Power System Alarm Data

Ricardo Bessa

INESCTEC

ricardo.j.bessa@inesctec.pt

Motivation and Challenges

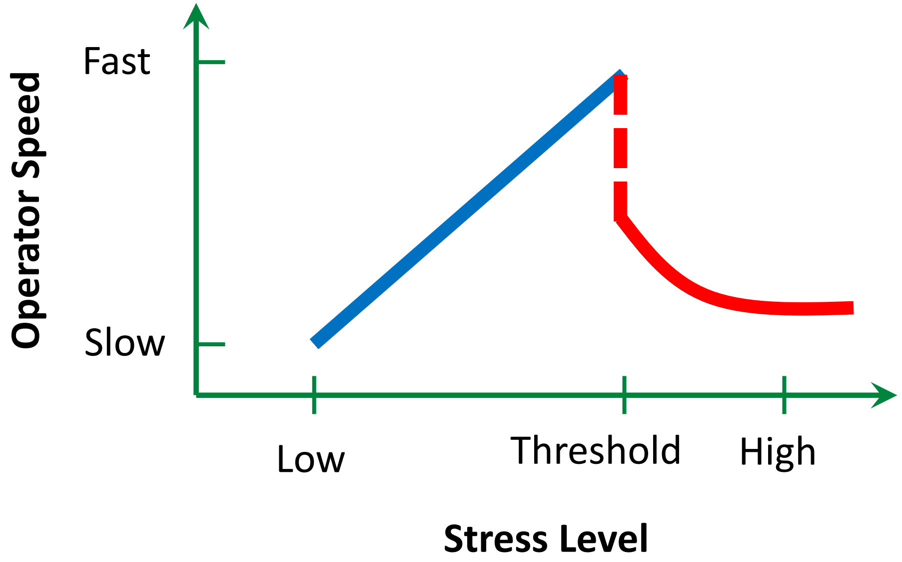


Fig. ref.: Kirschen, D. S., Wollenberg, B. F. (1992). Intelligent alarm processing in power systems. Proceedings of the IEEE, 80(5), 663-672

Challenges

Power system more complex to operate

Operation complexity rising at a fast pace
(e.g., DER, prosumers, grid maintenance)

Decision-making under pressure

Grid monitoring & protection equipment

Large volume of real-time and historical alarms data

High-volume of the information generated during outages

up to 200 alarm messages per minute

High Cognitive Load!

Alarms!

Decisions!

Data!

only for assets in HV and MV grids!

Data Representation

Raw SCADA Data from a DSO

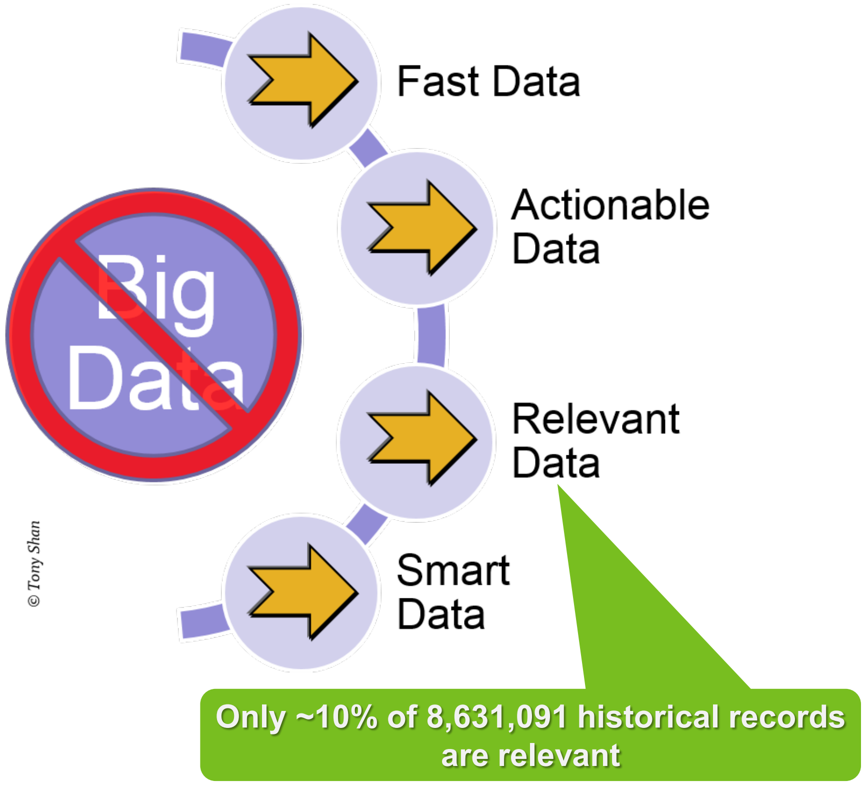


Fig. ref.: http://www.socialmediatoday.com/technology-data/2015-04-04/big-data-really-dead

evdate	evdesc
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I> INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P330 TRANSFORMADOR2 MAX I> INST UP1 ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P330 TRANSFORMADOR2 MIN U< INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P330 TRANSFORMADOR2 MIN U<< INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P502 TRANSFORMADOR2 MAX I> INST UP1 ARRANQUE
2014-01-02 06:33:13.000	SE SAO JORGE P502 TRANSFORMADOR2 MAX I> INST UP2 - DIF ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I>>> INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I>>> TEMP DISPARO
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE PROT DEFEITO FASE-FASE DISPARO
2014-01-02 06:33:24.000	SE SAO JORGE P332 SAO MAMEDE SUPERVISAO CIRCUIT DESL ALARME
2014-01-02 06:33:14.000	SE SAO JORGE P330 TRANSFORMADOR2 NORMALIZACAO TENSAO+FREQ INACTIVO
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE DISJUNTOR DESLIGADO
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE PROT TERRAS RESIST INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE MAX I>> INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE PROT TERRAS RESIST INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P329 TSA+REACTANCIA2 MAX Io> DTR INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P329 TSA+REACTANCIA2 MAX Io>>DTR INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I>>> INST NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE MAX I>>> INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE MAX I> INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P328 PATAIAS PROT TERRAS RESIST INST ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P329 TSA+REACTANCIA2 MAX Io>>>INST PHB ARRANQUE
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I> INST NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I>> INST NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE PROT DEFEITO FASE-FASE NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE PROT TERRAS RESIST INST NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P332 SAO MAMEDE MAX I>>> TEMP NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P328 PATAIAS PROT TERRAS RESIST INST NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE MAX I>>> TEMP DISPARO
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE PROT DEFEITO FASE-FASE DISPARO
2014-01-02 06:33:14.000	SE SAO JORGE P509 BARRAMENTO2 TENSAO BARR 56.386 KV Baixo
2014-01-02 06:33:14.000	SE SAO JORGE P329 TSA+REACTANCIA2 MAX Io>>>INST PHB NORMAL
2014-01-02 06:33:14.000	SE SAO JORGE P326 MIRA D'AIRE DISJUNTOR DESLIGADO

Source: E-REDES SCADA Alarm event log data
(a snapshot for less than a second)

Data Representation

Pros:
- simple approach
- low computational effort

Cons:
- unable to capture relations between sequence tags
- variable output length
- sparse representation (*one-hot encoding*)

JSRAN-3305-PRETA	ARRANQUE
JSRAN-3305-PRETA	ARRANQUE
JSRAN-3305-PRETD	DISPARO
JSRAN-3305-DJDT-	DISPARO
JSRAN-3305-DJEST	DESLIGADO

Pros:
- simple approach
- low computational effort
- fixed output length

Cons:
- unable to capture relations between sequence tags
- sparse representation of original data

Integer Encoding

JSRAN-3305-PRETA	ARRANQUE	1
JSRAN-3305-PRETD	DISPARO	2
JSRAN-3305-DJDT-	DISPARO	3
JSRAN-3305-DJEST	DESLIGADO	4

One-Hot Encoding

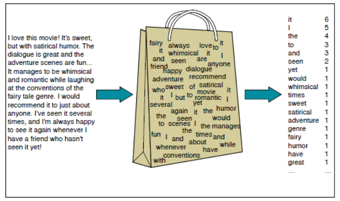
Human-Readable

Pet
Cat
Dog
Turtle
Fish
Cat

Machine-Readable

Cat	Dog	Turtle	Fish
1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1
1	0	0	0

Bag-of-Words



TF-IDF

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right)$$

tf_{ij} = number of occurrences of i in j
df_i = number of documents containing i
N = total number of documents

Transformed Sequence

[1, 1, 2, 3, 4]

(1 row per sequence)

(Max. number = Vocabulary size)

[1, 0, 0, …, 0]
[1, 0, 0, …, 0]
[0, 1, 0, …, 0]
[0, 0, 1, …, 0]
[0, 0, 0, …, 1]

(1 row per sequence)

(nr. cols = vocabulary size)

Tag1, Tag2, Tag3, Tag4
[2, 1, 1, 1]

(1 row per sequence)

(nr. cols = vocabulary size)

[0.2, 0.1, 0, 0, …, 0.9, 0.2, 0]

(1 row per sequence)

(nr. elements = vocabulary size)

Reference: J.R. Andrade, R.J. Bessa, et al., "Data-driven anomaly detection and event log profiling of SCADA alarms," IEEE Access, vol. 10, pp. 73758-73773, 2022.

Data Representation

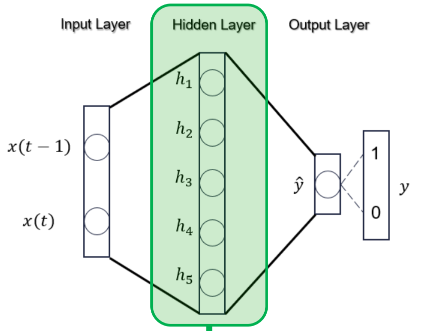
Pros:
- captures relations between sequence tags

Cons:
- time-consuming (parameter tunning needed)
- high computational effort
- variable output length

Word2Vec

Input Layer

Tag sequence skip-grams



Output

Probability of being the nearby tag for every tag in vocabulary

Word Embeddings Sequence:

[0.3, 0.2, 0.1, 0.9]	Tag1
[0.8, 0.5, 0.9, 0.3]	Tag2
[0.2, 0.3, 0.3, 0.1]	Tag3
[0.1, 0.7, 0.3, 0.2]	Tag4

N x M matrix

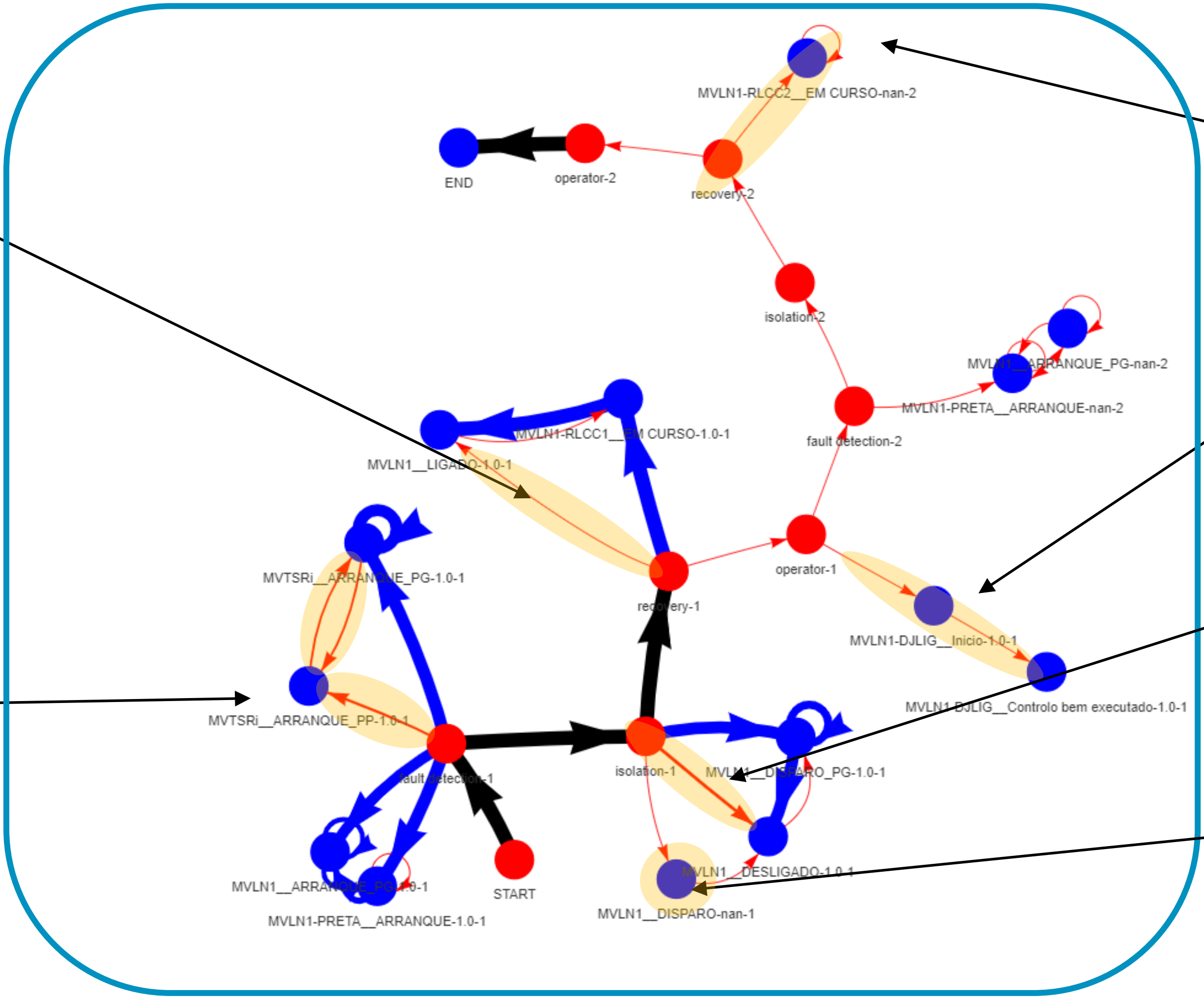
Where:
N – number of words in sequence
M – embedding layer dimension (n. units)

Reference: J.R. Andrade, R.J. Bessa, et al., "Data-driven anomaly detection and event log profiling of SCADA alarms," IEEE Access, vol. 10, pp. 73758-73773, 2022.

Graph-based Representation

Time synchronization problem
reclosure cycle event after circuit breaker closed

Abnormal pickup during phase-to-ground fault



Graph for a collection of short phase-to-ground cases

Operator sent command to close a circuit breaker after it was closed already

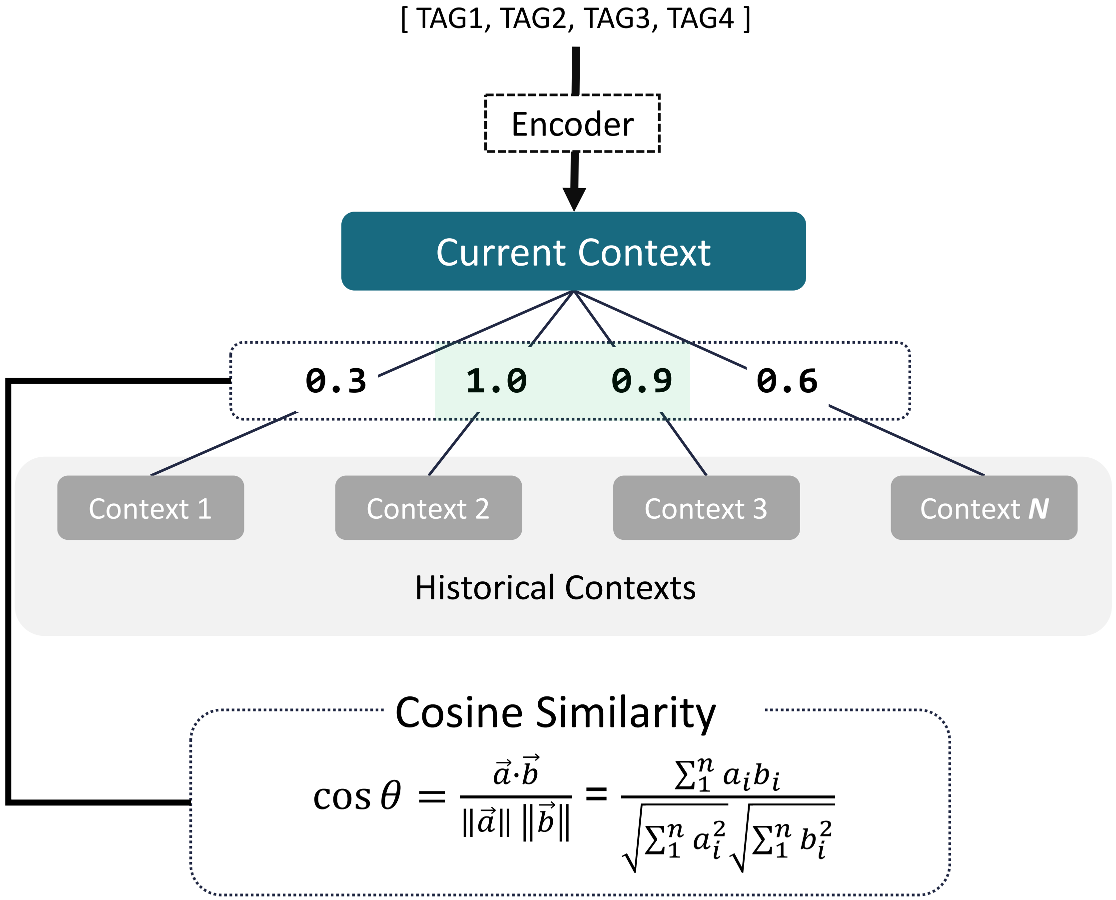
<u>Missing event</u>
Circuit breaker opens without a protection TRIP

<u>Reported as “Trabalho EDP Empreiteiro”</u>
“DISJ em REE DISPARO”
MV Line circuit breaker in “Special Exploration Regime” with automated reclosure cycles deactivated

Reference: V. Campos, O. Klyagina, J.R. Andrade, R.J. Bessa, C. Gouveia, "ML-Assistant for human operators to solve and classify fault occurrences in electrical grids," under review, 2023.

Example of Use Cases

UC1: Identification of Similar Events / Abnormal Protection Behaviour



Profile the "normal" protection behavior within the subset of contexts for the same fault type using data from all substations

Common tags	Uncommon tags	Rare tags

Detect missing "common" protection logs

(logs frequent in historical contexts but not present in current context)

Detect "uncommon" and "rare" protection logs

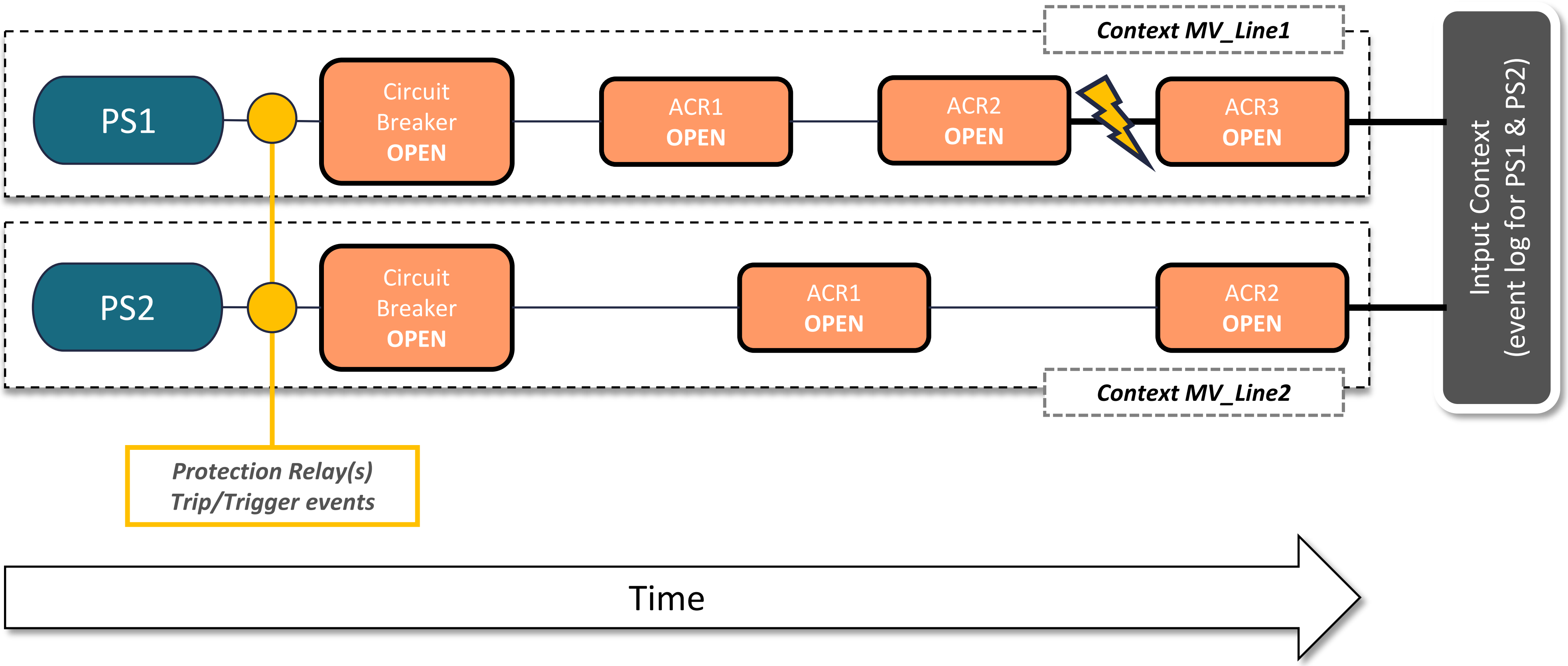
(logs present in current context and rarely detected in similar historical contexts)

Reference: J.R. Andrade, R.J. Bessa, et al., "Data-driven anomaly detection and event log profiling of SCADA alarms," IEEE Access, vol. 10, pp. 73758-73773, 2022.

UC2: Inference of the next operator commands

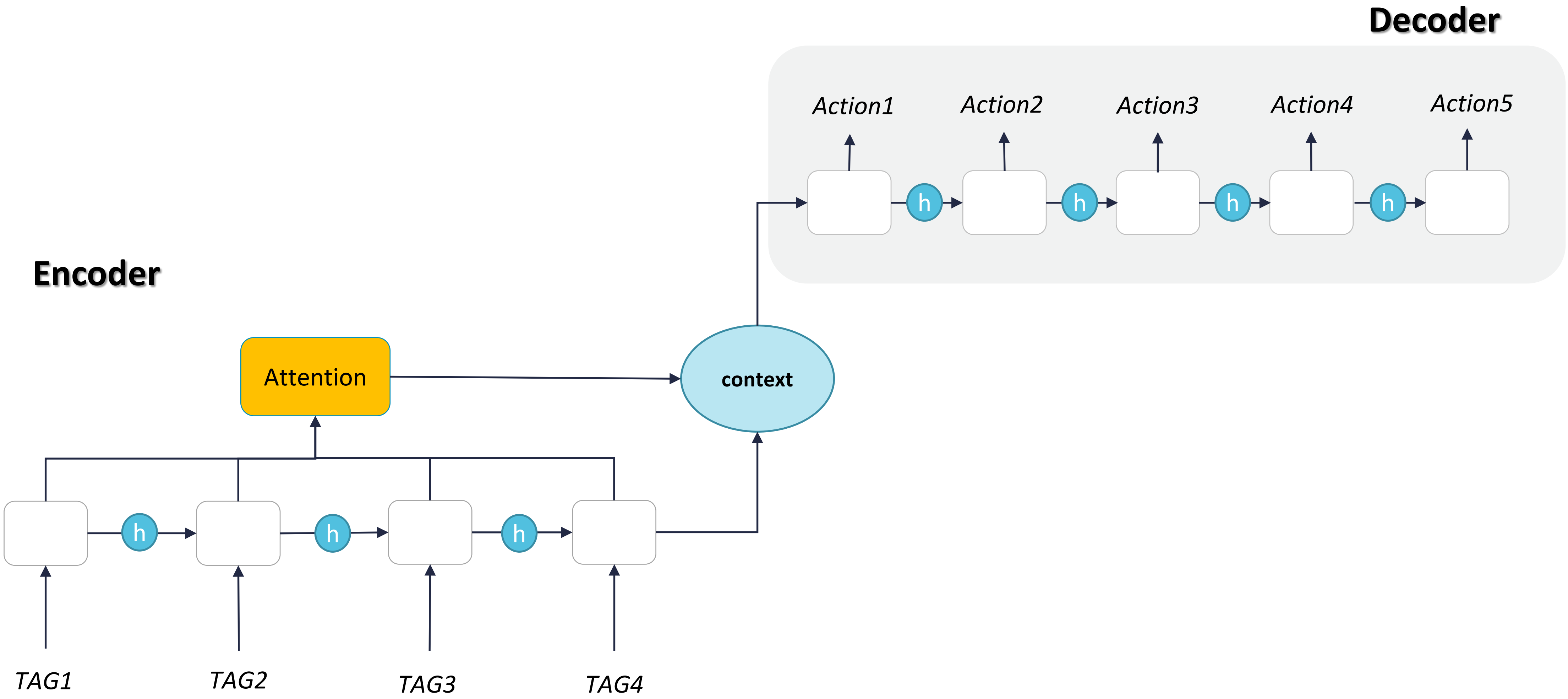
Occurrence

1. Fault in line segment between Automatic Circuit Reclosers (ACR) 2 and 3 protection equipment's
2. Circuit breaker opening for Primary Substation 1 and Primary Substation 2



Reference: V. Campos, R.J. Bessa, et al., "ML-assistant for human operators to solve faults and classify events complexity in electrical grids," MEDPOWER 2022.

UC2: Inference of the next operator commands

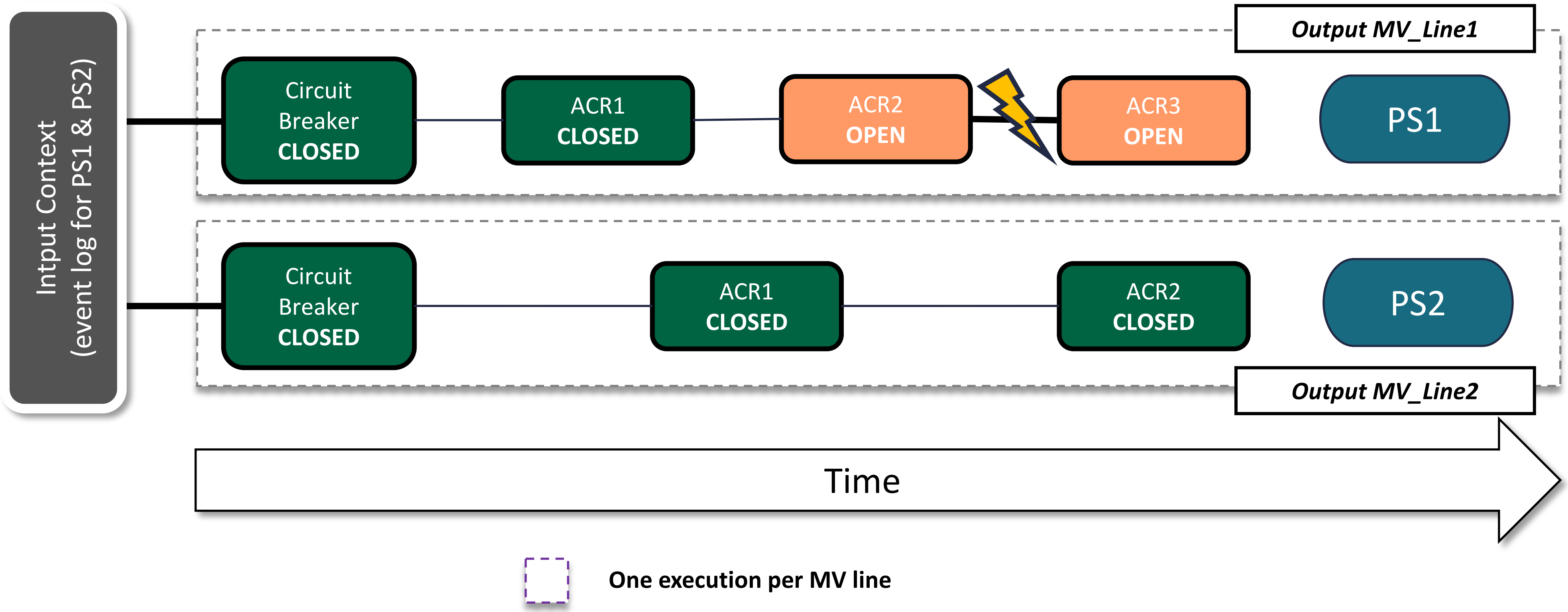


Reference: V. Campos, R.J. Bessa, et al., "ML-assistant for human operators to solve faults and classify events complexity in electrical grids," MEDPOWER 2022.

UC2: Inference of the next operator commands

Occurrence

1. Fault in line segment between ACR2 and ACR3 protection equipment's
2. Circuit breaker opening for Primary Substation 1 and Primary Substation 2
3. Isolation of affected segment of MV Line1 by leaving ACR2 and ACR3 in OPEN state
4. Service was completely restored in MV Line2 by closing the circuit breaker and all the ACRs throughout the line



Concluding Remarks

- Alarm data has been undermined…
  - …in value for system operation and event analysis, but also in human-computer interaction
- Data scarcity is a challenge: poor performance of data-intensive methods (deep learning)
- The potential of use cases for alarm data is high. Other examples explored at INESC TEC:
  - Classify type occurrences (simple ones can be replaced by automation)
  - Event log profiling and segmentation
  - Improve fault cause classification

Acknowledgments

Ricardo Andrade, Vasco Campos, Olga Klyagina
INESC TEC