ICDE 2023
Anaheim, CA April 3 – 7, 2023

Tele-Knowledge Pre-training for Fault Analysis

Zhuo Chen[1,4†], Wen Zhang[1,4†], Yufeng Huang[1], Mingyang Chen[1], Yuxia Geng[1], Hongtao Yu[1], Zhen Bi[1], Yichi Zhang[1], Zhen Yao[1], Wenting Song[2], Xinliang Wu[2], Yi Yang[2], Mingyi Chen[2], Zhaoyang Lian[2], Yingying Li[2], Lei Cheng[2], Huajun Chen[1,3,4‡]

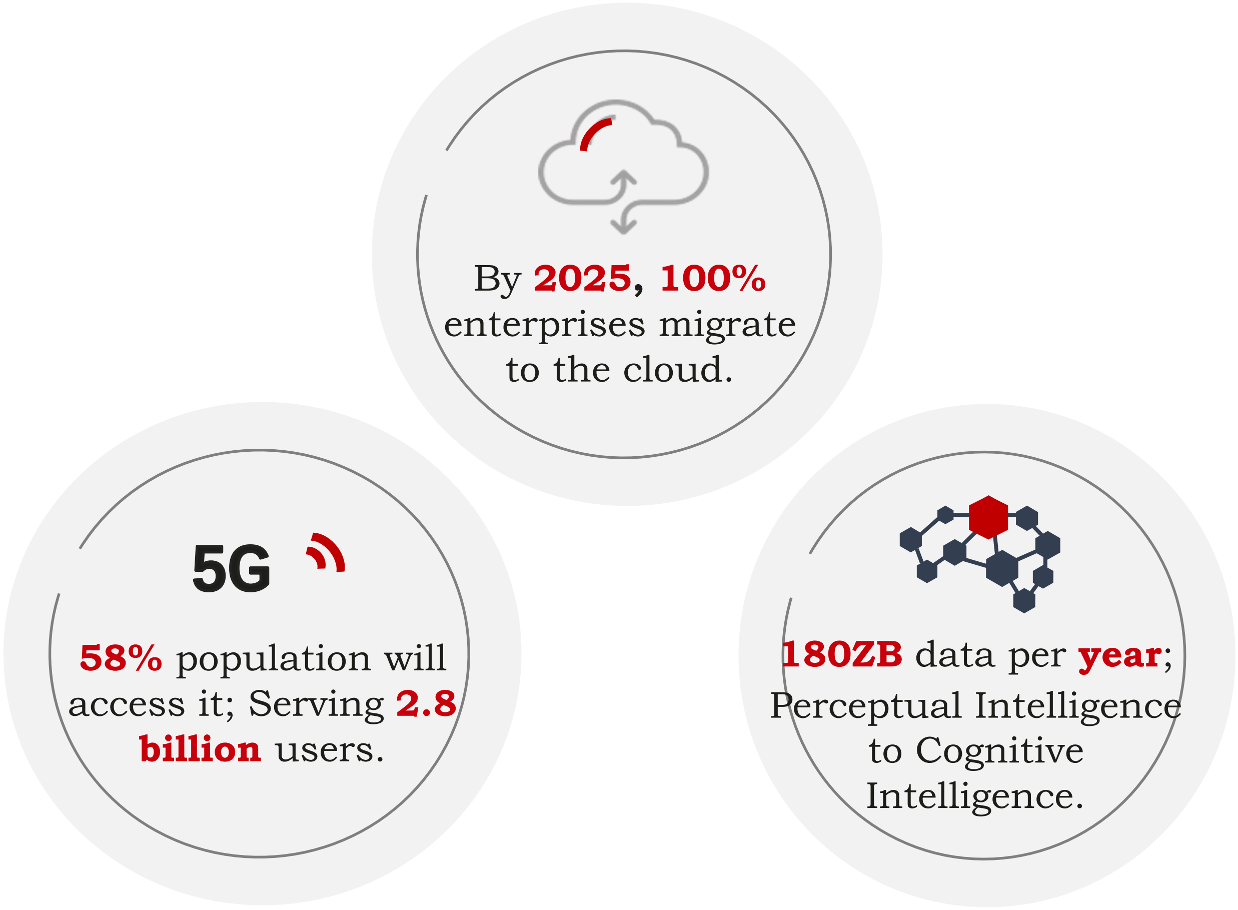
1 College of Computer Science and Technology, Zhejiang University, Hangzhou, China
2 NAIE PDU, Huawei Technologies Co., Ltd., Xi'an, China
3 Donghai Laboratory, Zhoushan 316021, China
4 Alibaba-Zhejiang University Joint Institute of Frontier Technologies, Hangzhou, China

浙江大學 ZHEJIANG UNIVERSITY

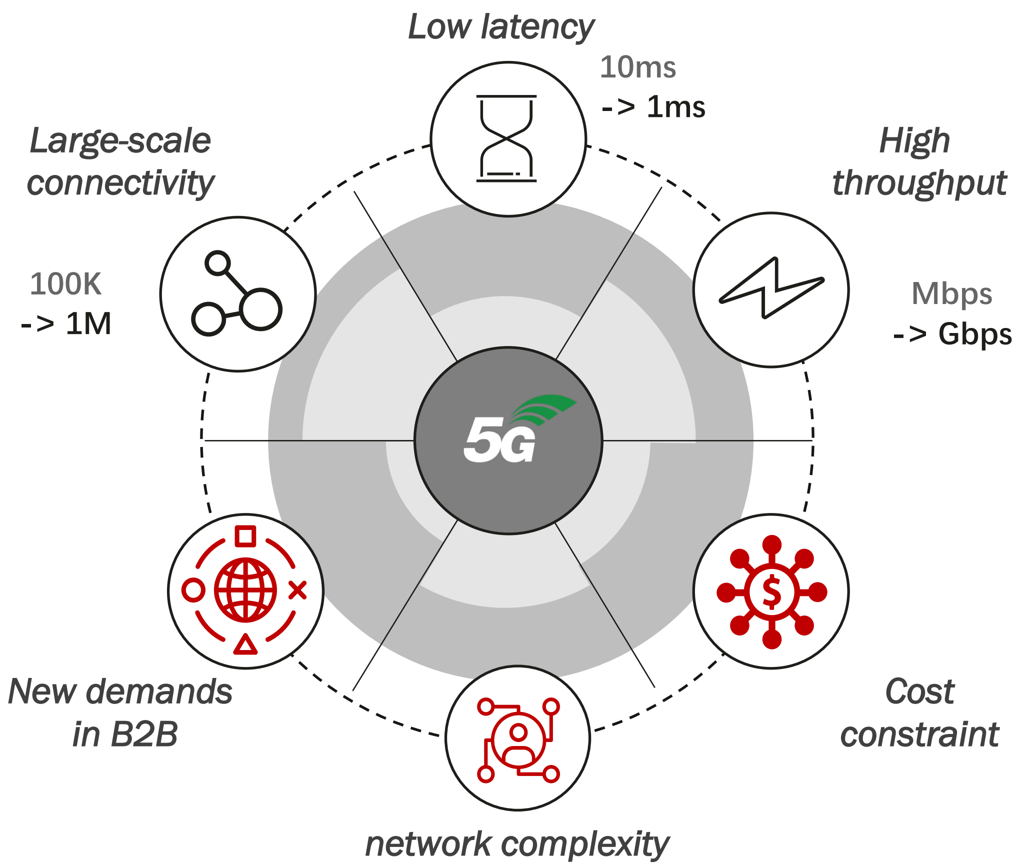
HUAWEI

Industry background:

◆**Multi-agent collaboration enhance the value of AI systems**



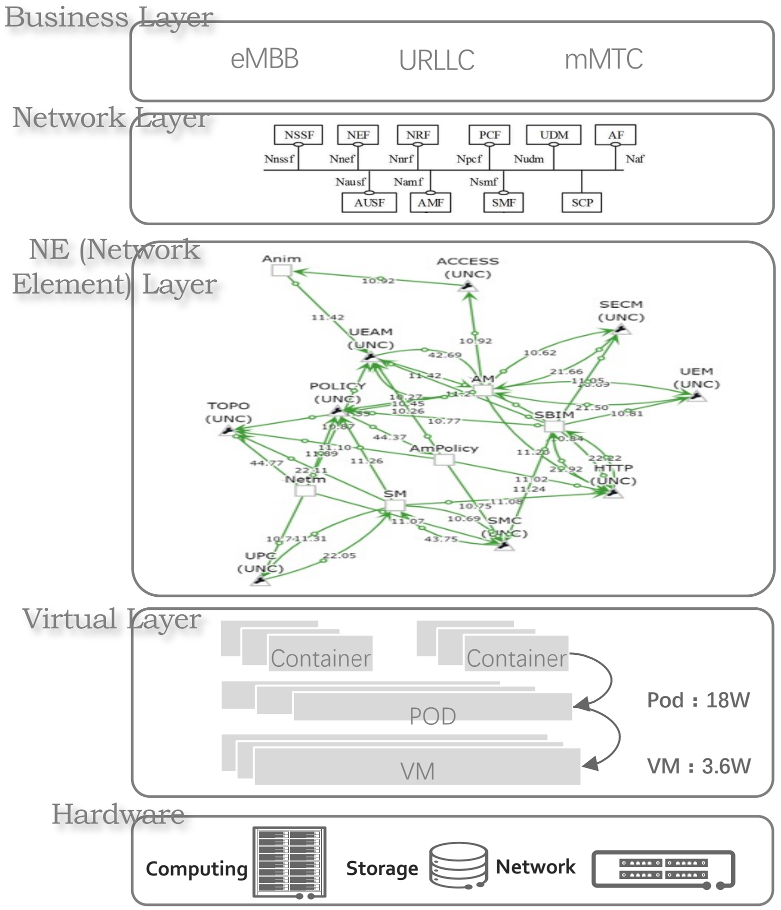
◆**New challenges call for highly intelligent and autonomous tele-networks.**



Industry challenges:

ICDE2023
Anaheim, CA April 3 – 7, 2023

◆**Cloud-based tele-networks pose challenges for traditional fault tree analysis**



Dynamic Changes of Business Slicing

Dynamic Changes of Network Topology

Dynamic Changes of Microservice Relations

Dynamic Changes of Resource Topology

Multi- vendor

- **High Number of Alarms**:
  - The core network has 7,000+ alarms and 230+ alarm-related work orders per day.
- **Various fault types**:
  - 7 major categories such as HTTP link failure, and thousands of subcategories.
- **Difficult to Locate**:
  - After 5G cloudification, locating faults in network element links is difficult
- **Difficult to Accumulate Experience**:
  - Fault analysis tools based on expert experience rules have a backlog in fault processing (unable to handle new faults that have never occurred before)

Corpus overview

Time	NE	Initial registration request number	Successful initial registration number	Initial registration failure number	Average duration of initial successful registration	Total duration of initial successful registration	Initial registration success rate
11:35:00	AMF1/SE	0.017985	0.030351	0	0.002289	0.000277	0.998962
11:40:00	AMF1/SE	0.018068	0.030483	0	0.002251	0.000277	0.998541
11:45:00	AMF1/SE	0.017755	0.029963	0	0.002209	0.00027	0.998962
11:50:00	AMF1/SE	0.017074	0.02879	0.5	0.002287	0.000261	0.997627
11:55:00	AMF1/SE	0.016259	0.027439	0	0.002307	0.000248	0.998962
12:00:00	AMF1/SE	0.016434	0.027726	0.25	0.002319	0.000251	0.998492
12:05:00	AMF1/SE	0.016779	0.028316	0	0.002363	0.000259	0.998962
12:10:00	AMF1/SE	0.016558	0.027936	0	0.002398	0.000257	0.998517

The numerical KPI data

Occur time	Time split	NE	Occur time_utc	Anormal value	Exception status	KPI name	KPI id	NF name
11:50:00	5	AMF1/SE	1647316200	393	Down	laximum user number of AMF idle state	1.929E+09	AMF1
11:30:00	5	AMF1/SE	1647315000	12010	Up	imum user number in the AMF connecti	1.929E+09	AMF1
11:10:00	5	AMF2/SE	1647313800	232041	Up	discovery request number initiated w	1.929E+09	AMF2
11:30:00	30	AMF2/SE	1647315000	74932	Up	authentication request number in N2	1.929E+09	AMF2
11:35:00	5	AMF2/SE	1647315300	1922	Up	number in N2 mode authentication pa	1.929E+09	AMF2
11:50:00	5	AMF2/SE	1647316200	189654	Down	3GPP deregration number initiated by	1.929E+09	AMF2
11:35:00	5	AMF2/SE	1647315300	20750	Up	duration of initial successful regis	1.929E+09	AMF2
11:35:00	5	AMF2/SE	1647315300	3E+09	Up	luration of initial successful regist	1.929E+09	AMF2
11:35:00	5	AMF2/SE	1647315300	59.59	Down	Initial registration success rate	1.929E+09	AMF2
11:30:00	30	AMF2/SE	1647315000	31313	Up	mber of AMF registered real-time use	1.929E+09	AMF2
11:50:00	5	AMF1/SE	1647275100	99.946	Down	Initial registration success rate	1.929E+09	AMF1
11:30:00	30	AMF2/SE	1647315000	31313	Up	umber of AMF connected real-time user	1.929E+09	AMF2
11:30:00	30	AMF2/SE	1647315000	34231	Up	verage number of AMF registered user	1.929E+09	AMF2
11:35:00	5	AMF2/SE	1647315300	548	Up	verage number of AMF idle state user	1.929E+09	AMF2
11:30:00	30	AMF2/SE	1647315000	34231	Up	rage number of AMF connected state us	1.929E+09	AMF2
11:30:00	30	AMF2/SE	1647315000	49195	Up	umber of users in the AMF registrati	1.929E+09	AMF2
11:35:00	5	APP-HXB	1647315300	451285	Up	Total number of responses	1.931E+09	UDM001
11:35:00	5	APP-HXB	1647315300	1922	Up	Total number of paired-end errors	1.931E+09	UDM001

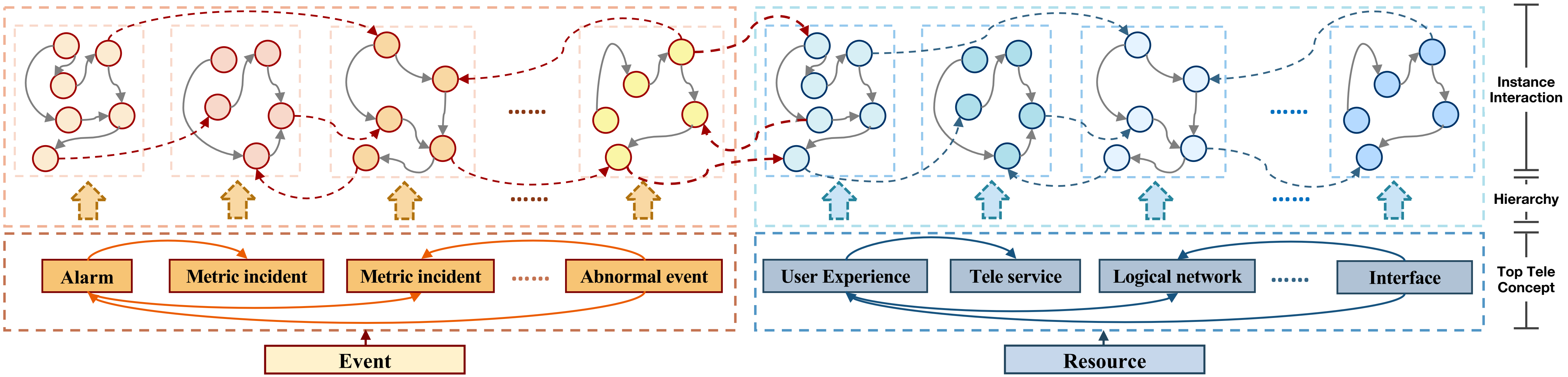
Abnormal event

(*a*) Machine (Log) Data.

ALM-81011 SIG Knowledge base upgrade failed

Explanation: 1) Alarm trigger mechanism: The system will generate this alarm when the SIG knowledge base upgrade fails. Then the system will continue to run according to the previously successfully loaded version of the knowledge base, so that recognition ability before the upgrade will not be affected;
2) Alarm recovery mechanism: ...
Attribute: Alarm_ID: 81011; Alarm_Level: Importance; Automatically cleared: Yes
Parameter: POD name: ... ; NE name: ... ; Event type: ...
Impact on the system: 1) Protocols and adaptation relations defined in the new knowledge base are not available. 2) ...
Possible reason: 1) The knowledge base digital signature file does not exist; Failure due to internal processing error; 2) ...

(*b*) Product Documents.

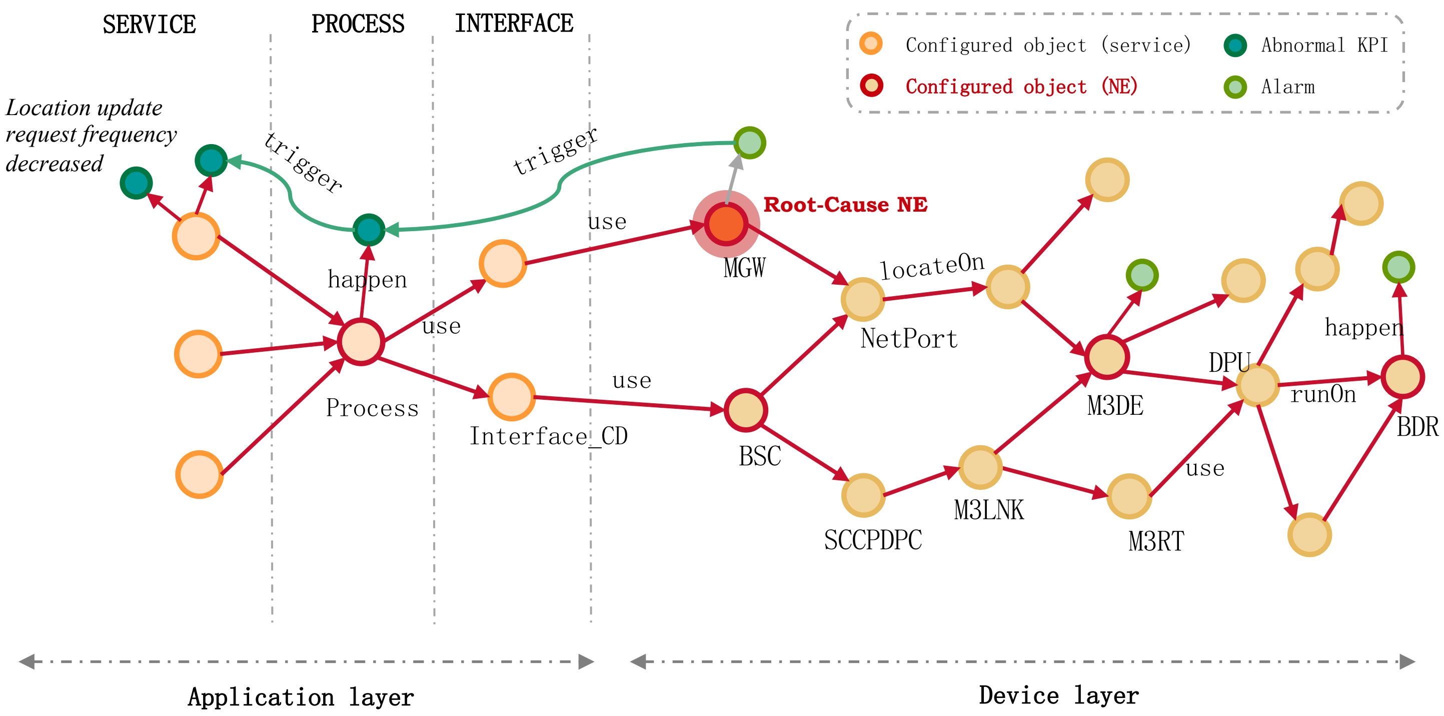


(*c*) Tele-product Knowledge Graph (Tele-KG).

Task abstraction: (1/4)

ICDE2023
Anaheim, CA April 3 – 7, 2023

- **Root-Cause Analysis:**
  - Identify the network element (NE) that is most likely to be the source of a fault in a tele-network.
- **Event Association Prediction:**
  - Automatically predicting the trigger relationship in candidate event pair
- **Fault Chain Tracing**
  - Tracing the source of these failures in network when alarms were raised
- **Abnormal KPI Detection**
  - Discover the abnormal value change within the real-time KPI flow of the tele-system.



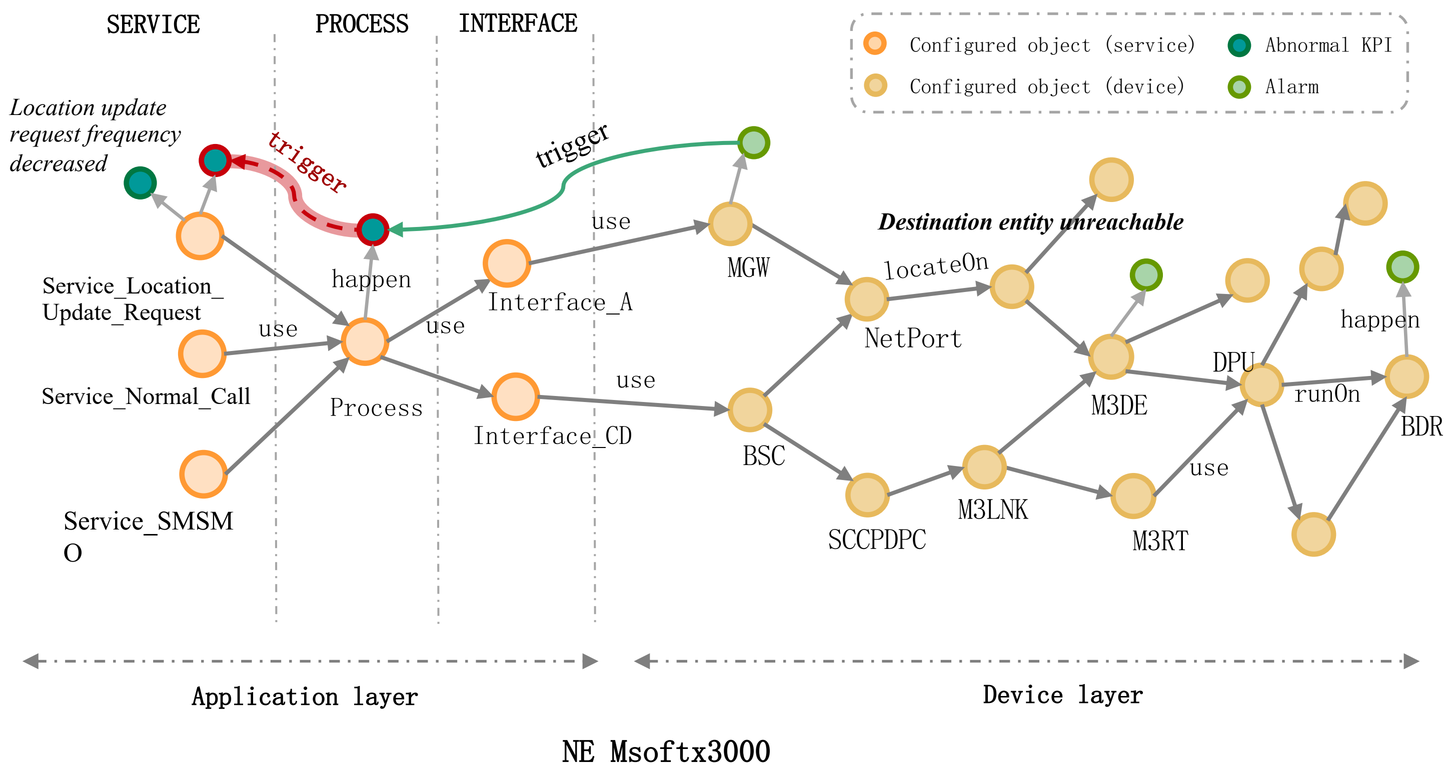
Task abstraction: (2/4)

- **Root-Cause Analysis:**
  - Identify the network element (NE) that is most likely to be the source of a fault in a tele-network.

- **Event Association Prediction:**
  - Automatically predicting the trigger relationship in candidate event pair

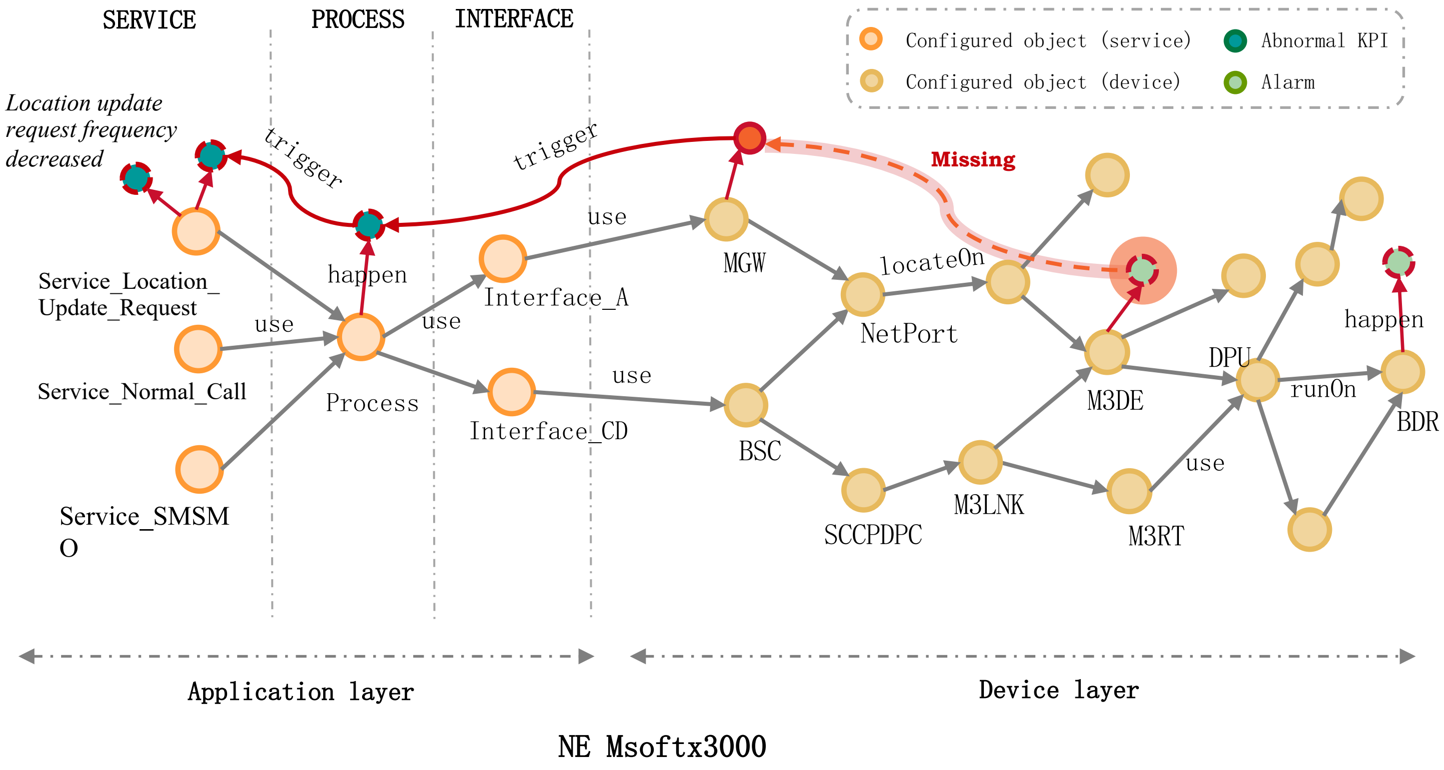
- **Fault Chain Tracing**
  - Tracing the source of these failures in network when alarms were raised

- **Abnormal KPI Detection**
  - Discover the abnormal value change within the real-time KPI flow of the tele-system.



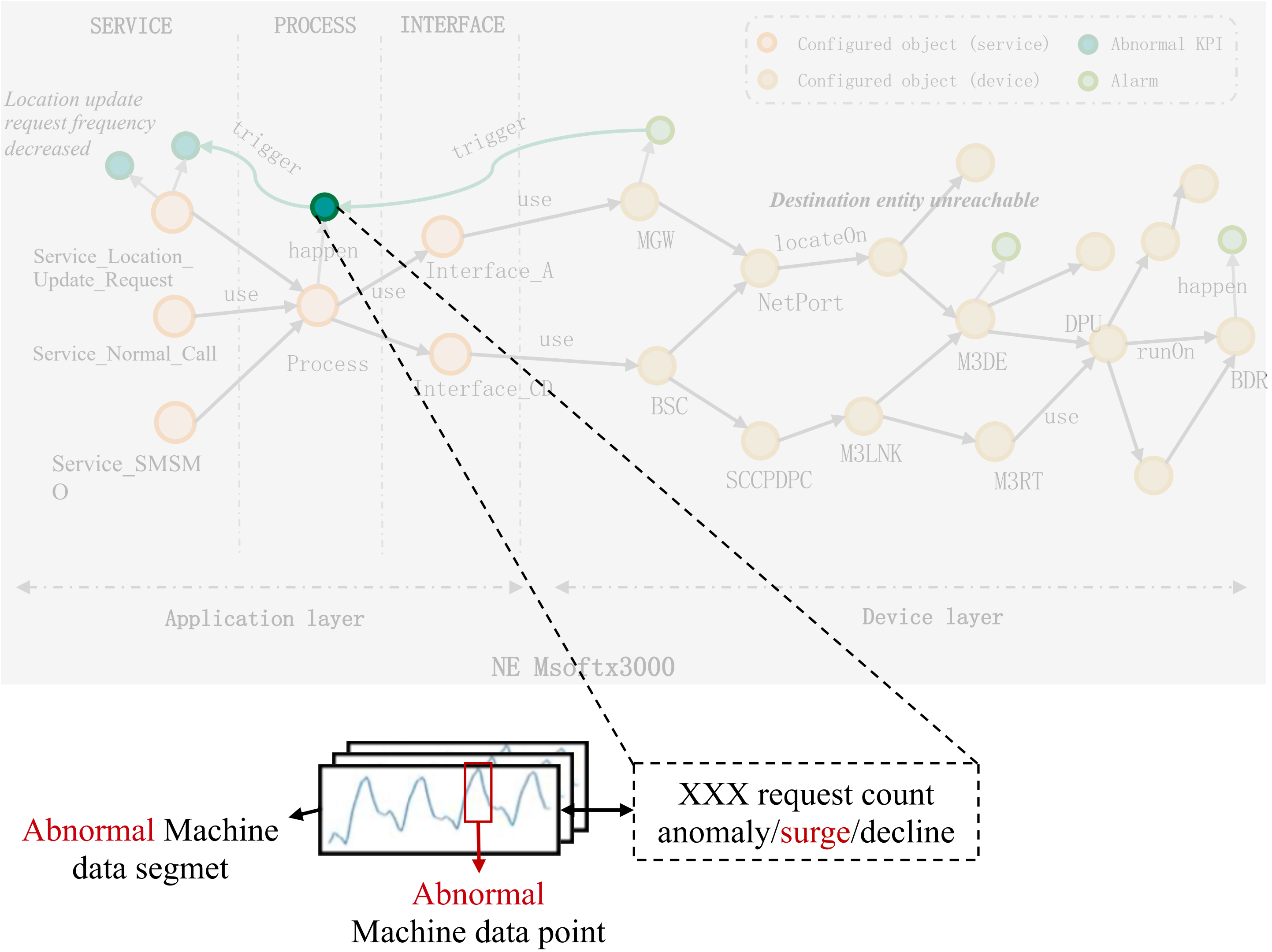
Task abstraction: (3/4)

- **Root-Cause Analysis:**
  - Identify the network element (NE) that is most likely to be the source of a fault in a tele-network.
- **Event Association Prediction:**
  - Automatically predicting the trigger relationship in candidate event pair
- **Fault Chain Tracing**
  - Tracing the source of these failures in network when alarms were raised
- **Abnormal KPI Detection**
  - Discover the abnormal value change within the real-time KPI flow of the tele-system.



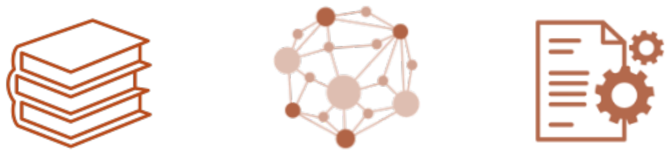
Task abstraction: (4/4)

- **Root-Cause Analysis:**
  - Identify the network element (NE) that is most likely to be the source of a fault in a tele-network.
- **Event Association Prediction:**
  - Automatically predicting the trigger relationship in candidate event pair
- **Fault Chain Tracing**
  - Tracing the source of these failures in network when alarms were raised
- **Abnormal KPI Detection**
  - Discover the abnormal value change within the real-time KPI flow of the tele-system.



Tele-domain Pre-training

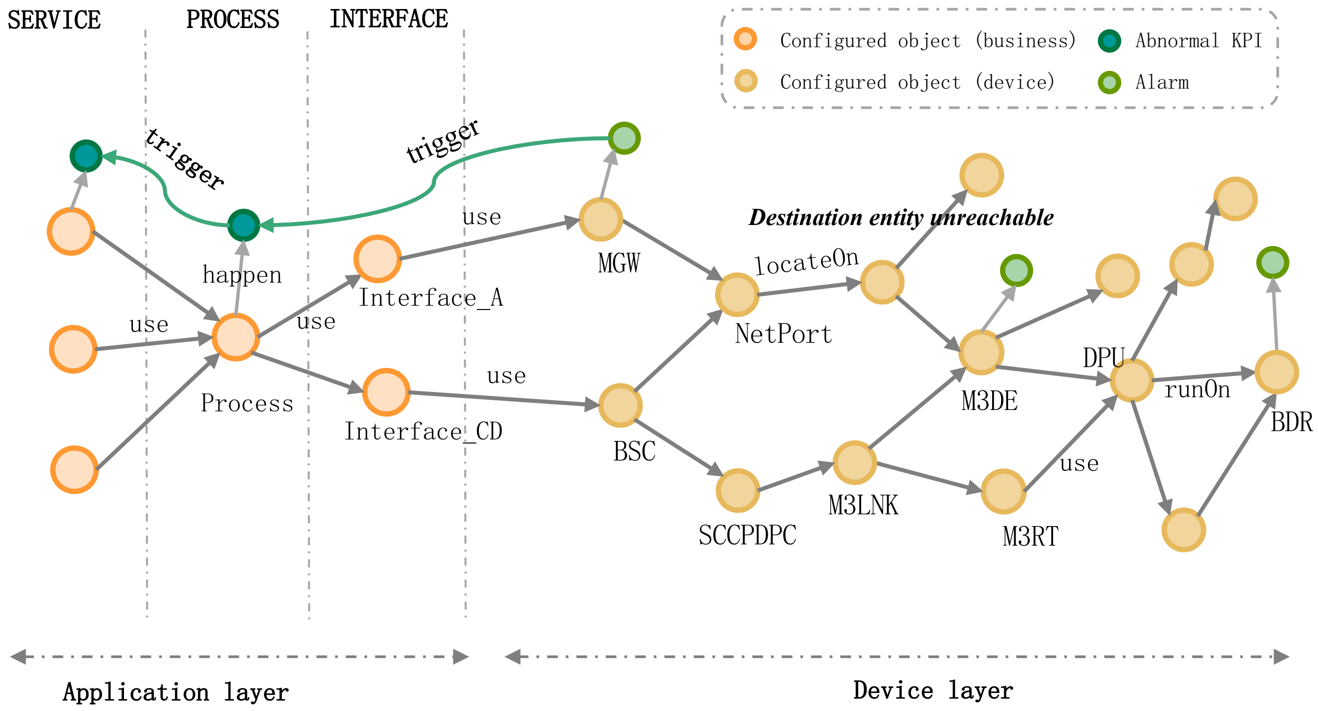
- **Technical Routes :**
  - **GPT** V.S. **BERT**✓
  - Decoder-based methods have limitations on Large-scale structure reasoning
- **Service Delivery Paradigm** :
  - Service embedding



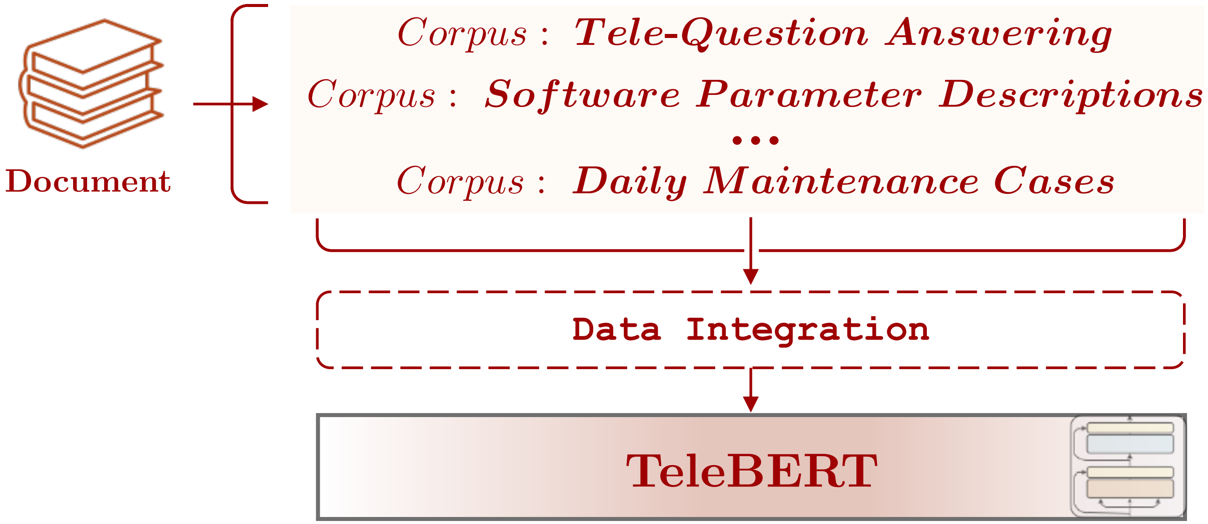
Document   Tele-KG   Machine Data

Aggregation Model

Task : ***Root-Cause Analysis***
Task : ***Event Association Prediction***
Task : ***...***
Task : ***Fault Chain Tracing***



TeleBERT：Tele-Document Pre-training



- *20,330 k sentences (1.4GB)*
- *8×8 32G NVIDIA V100 cluster*
- *30 Epoch*
- *4096 Batch size*
- *269 Hours*

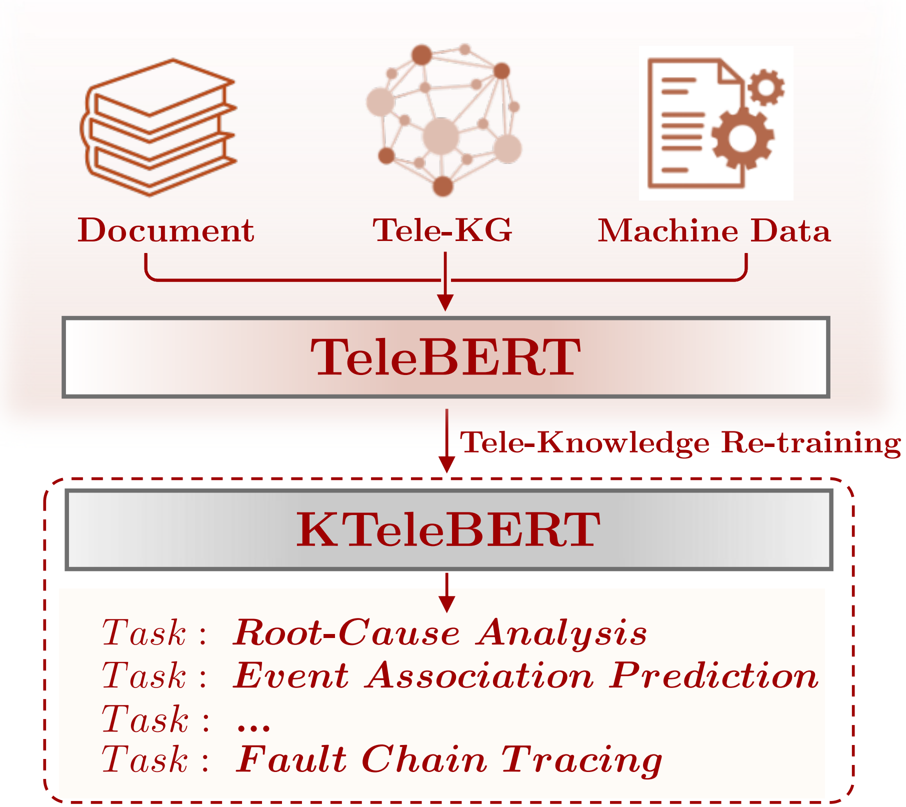
(a) Vanilla MLM training objective with ***whole word masking (WWM)*** *strategy*

(b) Chinese PLM MacBERT *as the backbone (≈372k elements)*

*(c) Contrastive learning on sentence embeddings (****SimCSE****)*

(d) The ***ELECTRA*** *pre-training paradigm*

KTeleBERT：Tele-Knowledge Re-training



- *434K causal sentences*
- *429K machine data*
- *130K knowledge triples*
- *4×24G NVIDIA RTX 3090*
- *60K steps*
- *256 Batch size*
- *8 Hours*

✓ *Unifying Modalities and Patterns*

✓ *Numerical Data Encoding*

✓ *Expert Knowledge Injection*

✓ *Multi-task learning for Multi-source Data*

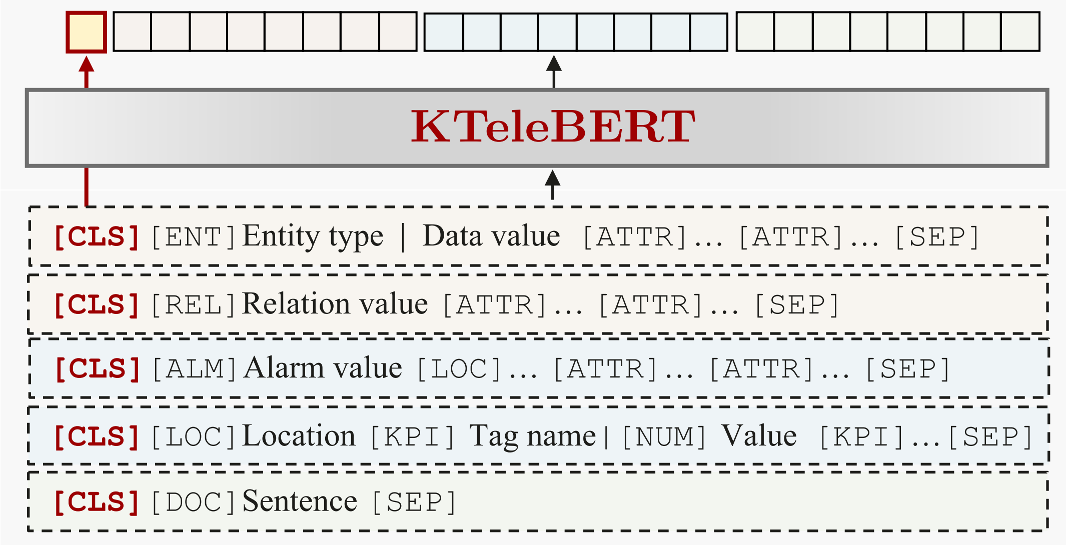
KTeleBERT：Tele-Knowledge Re-training

- **Unifying Modalities and Patterns** → *Implicit Knowledge Injection*

(a) *Causal sentences extraction (200k)*
  - *Causal meaning words/phrases*
  - *Extraction rules (e.g., minimum length)*
  - *Tele-KG Serialization*

(b) *Prompt template construction*

(c) *Tele special token construction*



KTeleBERT：Tele-Knowledge Re-training

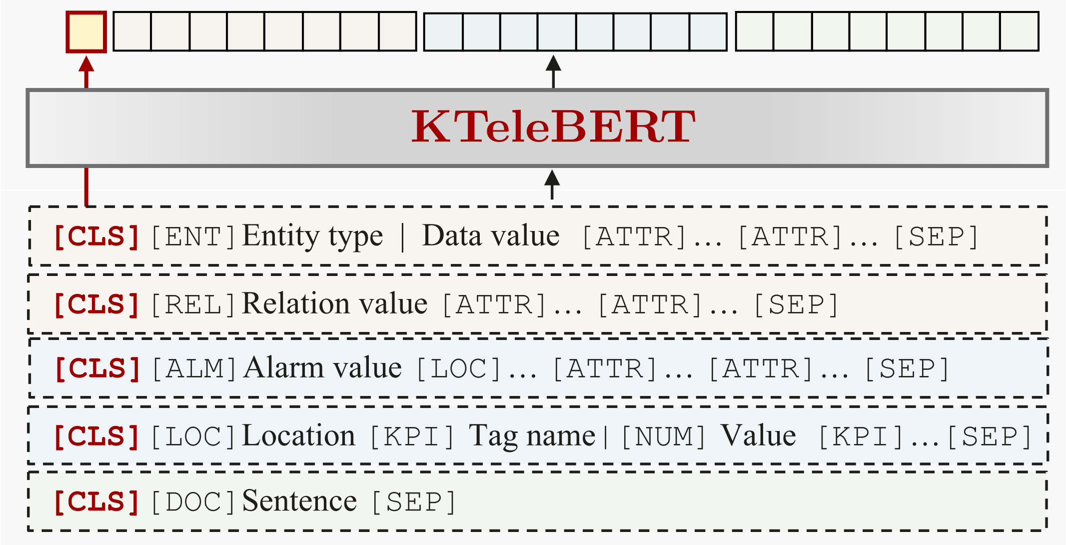
- **Unifying Modalities and Patterns** → *Implicit Knowledge Injection*

(a) Causal sentences extraction

(b) Prompt template construction (Instruction)

- *Prompts to represent the category*
- *Wrap the input with prompt templates*
- *Alleviates the disorder issue*

(c) Tele special token construction



KTeleBERT：Tele-Knowledge Re-training

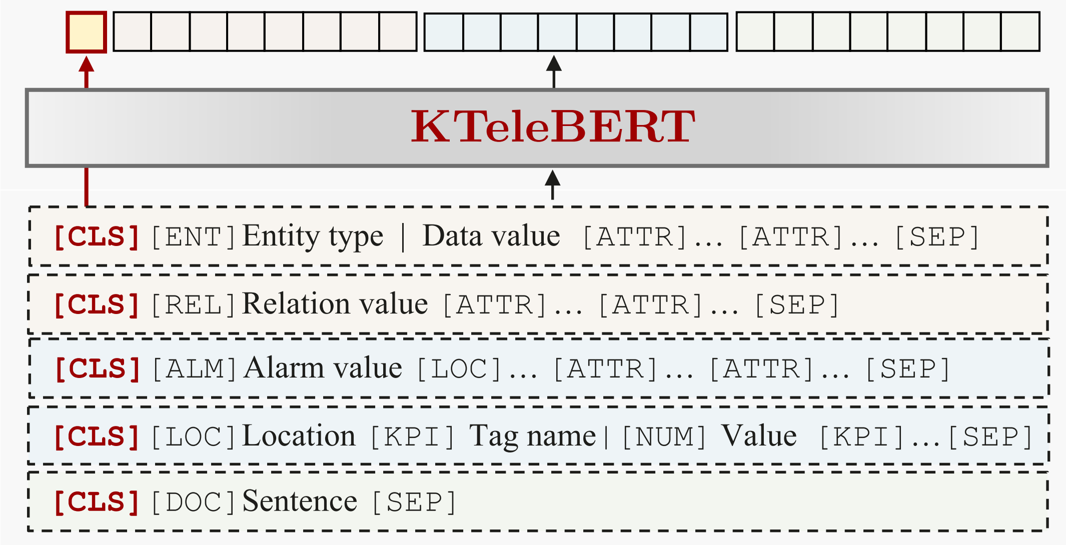
- **Unifying Modalities and Patterns** → *Implicit Knowledge Injection*

(a) *Causal sentences extraction*

(b) *Prompt template construction*

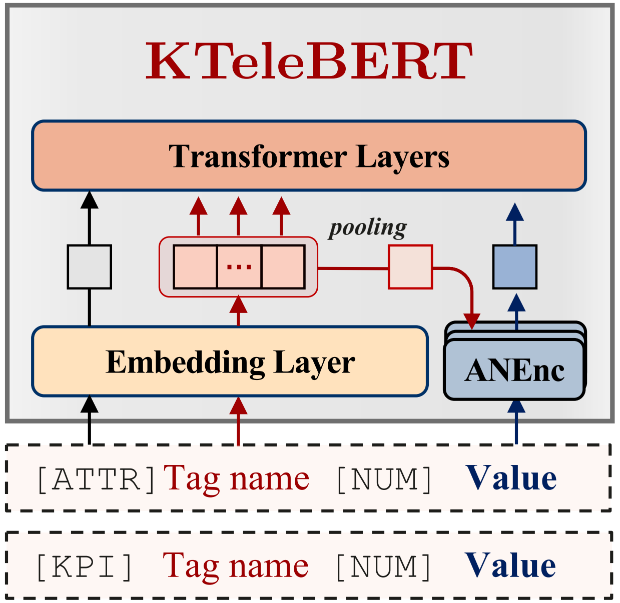
(c) ***Tele special token construction***

  - *Byte Pair Encoding (BPE)*
  - *domain-specific phrases or nouns*
    - *length: 2~4*
    - *Frequency > 8000*
    - *E.g., RAN, MML, MME, NF*



KTeleBERT：Tele-Knowledge Re-training

- **Numerical Data Encoding** → *Machine (Log) Data*



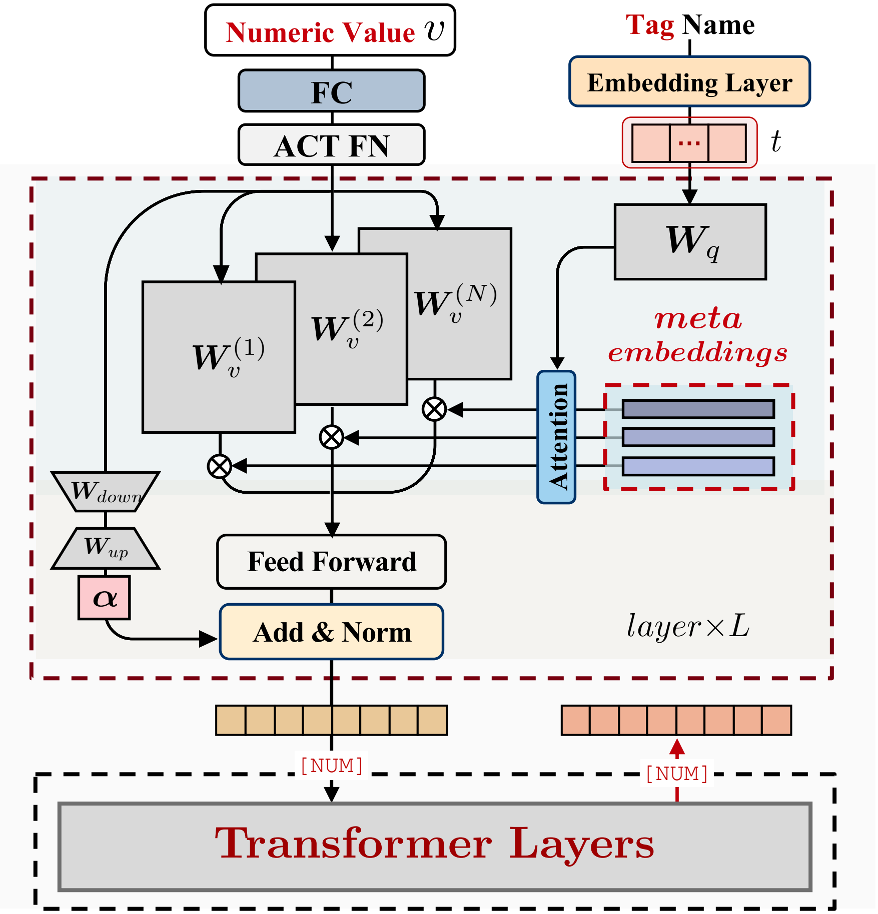
- *Information contained in Machine Data*
  - *Numerical value paired*
  - *Implied associations reflected in their correlated variations in value*
  - *A valuable supplement to expert experience in the tele-domain*

Time	NE	Initial registration request number	Successful initial registration number	Initial registration failure number	Average duration of initial successful registration	Total duration of initial successful registration	Initial registration success rate
11:35:00	AMF1/SE	0.017985	0.030351	0	0.002289	0.000277	0.998962
11:40:00	AMF1/SE	0.018068	0.030483	0	0.002251	0.000277	0.998541
11:45:00	AMF1/SE	0.017755	0.029963	0	0.002209	0.00027	0.998962
11:50:00	AMF1/SE	0.017074	0.02879	0.5	0.002287	0.000261	0.997627
11:55:00	AMF1/SE	0.016259	0.027439	0	0.002307	0.000248	0.998962
12:00:00	AMF1/SE	0.016434	0.027726	0.25	0.002319	0.000251	0.998492
12:05:00	AMF1/SE	0.016779	0.028316	0	0.002363	0.000259	0.998962
12:10:00	AMF1/SE	0.016558	0.027936	0	0.002398	0.000257	0.998517

KTeleBERT：Tele-Knowledge Re-training

- **Numerical Data Encoding** → *Machine (Log) Data*

- *ANEnc layer*
  - ➢ *Attention-based numeric projection (ANP)*
  - ➢ *Learnable field-aware meta embeddings*
    - *Decouple domain knowledge*
  - ➢ *Corresponding Value conversion function*
    - *Numerical embedding transformation*
  - ➢ *Query projection*
    - ➢ *Tag name embedding conversion*

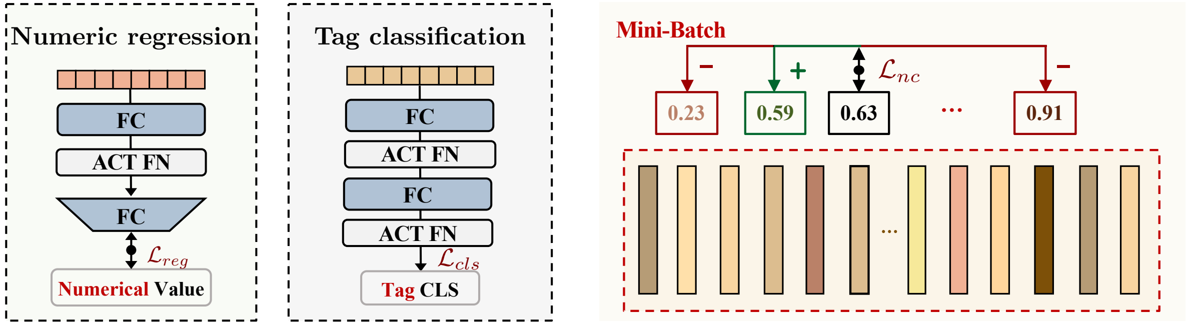


(a) Adaptive numeric encoder (ANEnc)

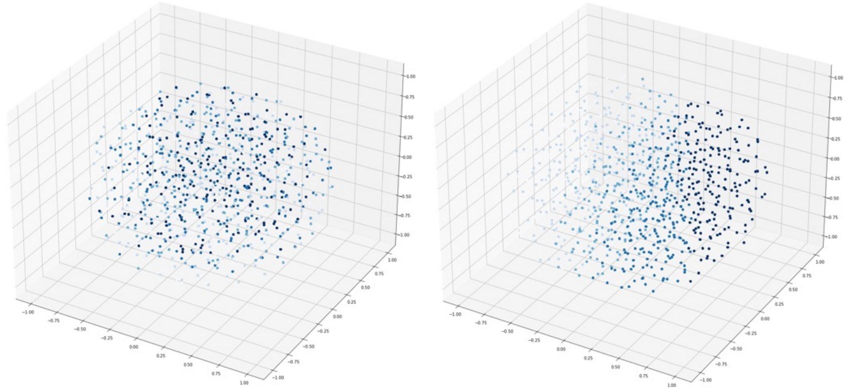
KTeleBERT：Tele-Knowledge Re-training

- **Numerical Data Encoding** → *Machine (Log) Data*

- *Training Objective*
  - ➢ *Numeric regression*
  - ➢ *Tag name classification*
  - ➢ *Numerical contrastive learning*
    - *Positive: closest value*
    - *Smooth numerical value changing* →
    - *Stabilize the model*
  - ➢ *Automatically weighted loss*
  - ➢ *Orthogonal regularization for Parameters*



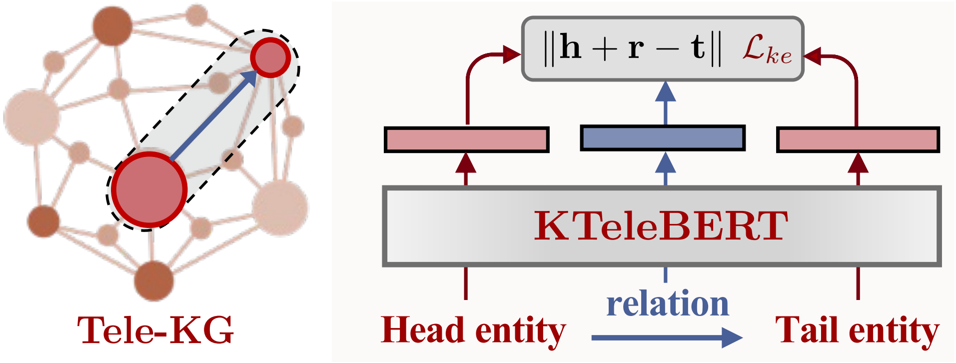
(b) Numeric decoder (c) Tag classifier (d) Numerical contrastive learning



(a) without $\mathcal{L}_{nc}$ (b) with $\mathcal{L}_{nc}$

- **Expert Knowledge Injection** → *Explicit Knowledge Injection*

- *Text-enhanced knowledge embedding (KE)*
  - *Following KEPLER*
  - *Injects knowledge into the model*
    - *Keep the original learning objectives*
    - *Maintain the original architecture*



KTeleBERT：Tele-Knowledge Re-training

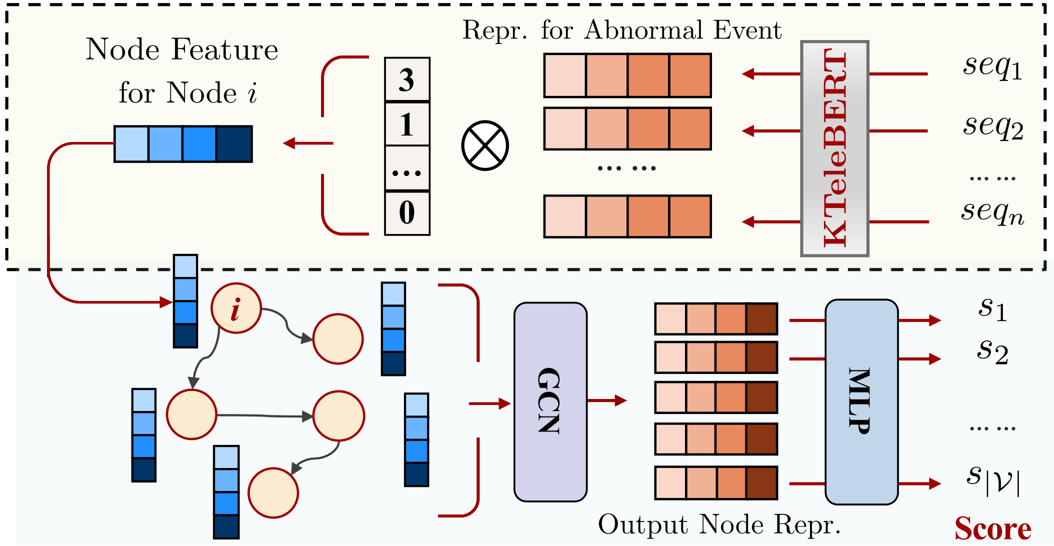
- **Multi-task learning for Multi-source Data**

  - *Multi-task Learning*
    - *Preventing Catastrophic Forgetting*
      - *Iterative multi-task learning (IMTL)*
      - *Cooperative parallel multi-task learning (PMTL)*

Strategy	Re-training task	Training iterations (steps)			Training objective
		Stage 1	Stage 2	Stage 3	
Single-task Learning (STL)	Masking Reconstruction		60k		$\mathcal{L}_{num} + \mathcal{L}_{mask}$
Parallel Multi-task Learning (PMTL)	Masking Reconstruction		60k		$\mathcal{L}_{num} + \mathcal{L}_{mask} + \mathcal{L}_{ke}$
	Knowledge Embedding		60k		
Iterative Multi-task Learning (IMTL)	Masking Reconstruction	40k	10k	10k	$\mathcal{L}_{num} + \mathcal{L}_{mask}$
	Knowledge Embedding	-	40k	20k	$\mathcal{L}_{ke}$

Task 1:

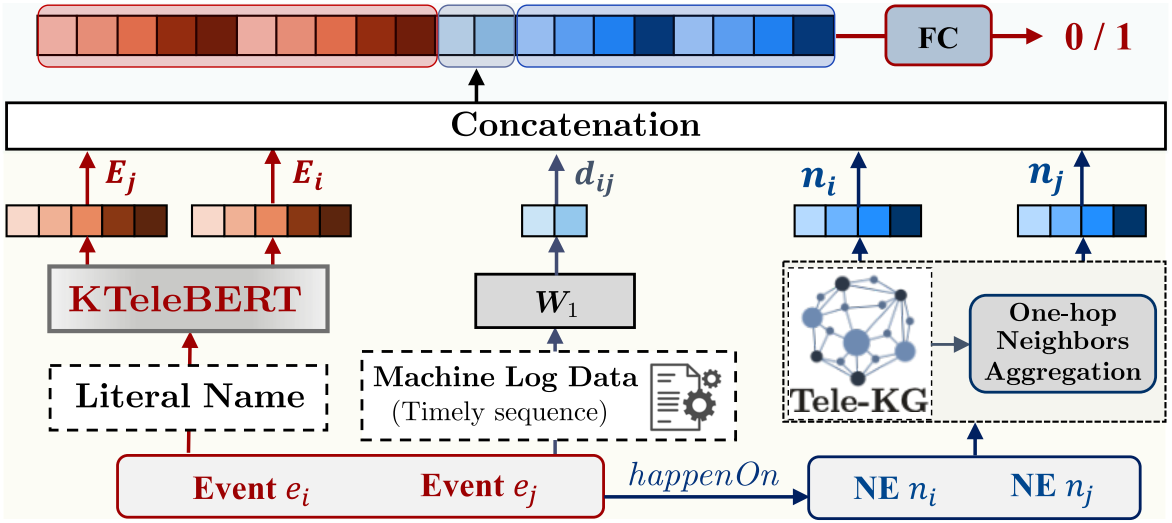
- *Root-Cause Analysis*
  - *Identify the network element (NE) that is most likely to be the source of a fault in a tele-network*
  - *Node ranking problem*



Method	MR ↓	Hits@1	Hits@3	Hits@5
Random	2.47	54.88	75.00	88.67
MacBERT	2.16	59.64	82.68	90.85
TeleBERT	2.09	62.65	83.52	92.46
KTeleBERT-STL	2.06	63.66	83.21	91.87
w/o ANEnc	2.13	60.72	82.96	90.80
KTeleBERT-PMTL	2.03	**65.96**	84.98	**92.63**
KTeleBERT-IMTL	**2.02**	64.78	**85.65**	91.13

Task 2:

- *Event Association Prediction*
  - *Representing events in low-dimensional embeddings (initialization)*
  - *Learning the associations between events*
  - *Trigger relation-specific space*

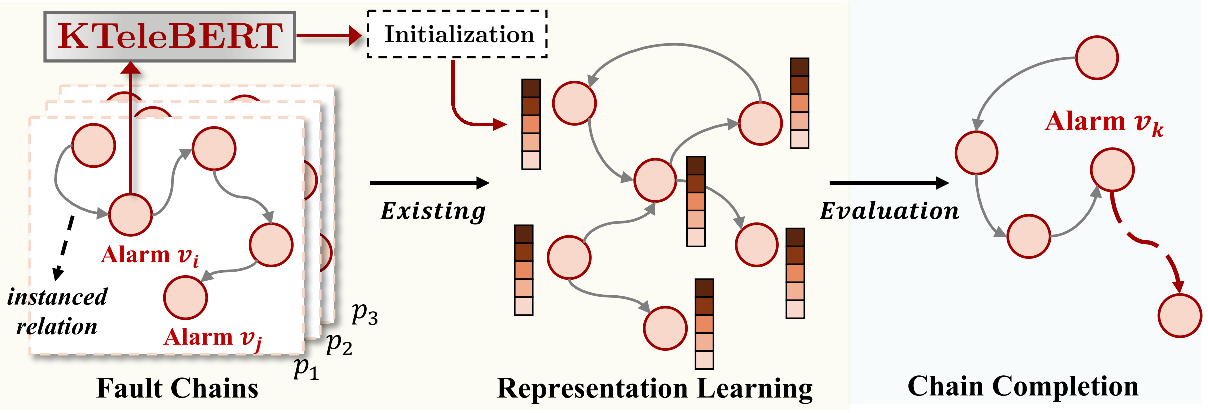


Methods	Accuracy	Precision	Recall	F1-score
Word Embeddings	64.9	66.4	96.8	78.7
MacBERT	64.3	65.9	96.1	78.2
TeleBERT	70.4	71.4	95.1	81.5
KTeleBERT-STL	**77.3**	**76.6**	96.6	**85.4**
w/o ANEnc	76.0	76.1	95.1	84.5
KTeleBERT-PMTL	68.5	68.8	**99.1**	81.3
KTeleBERT-IMTL	73.5	73.8	95.6	83.2

Task 3:

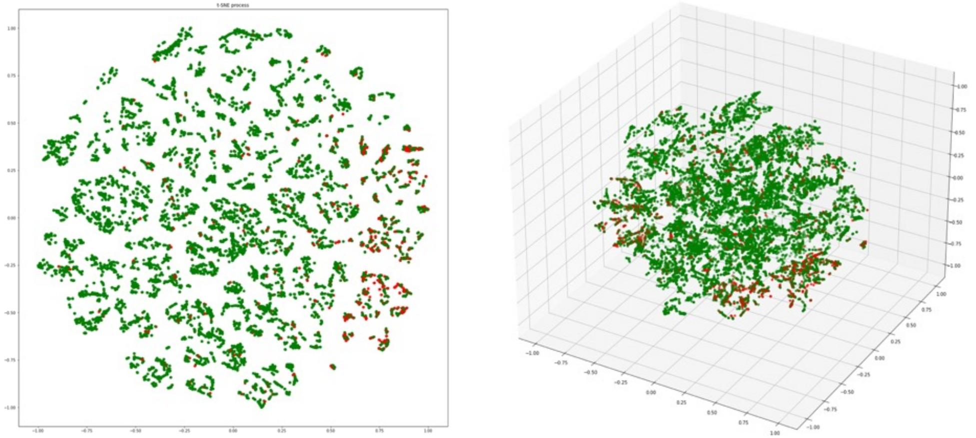
- ***Fault Chain Tracing***
  - *Connect those faults in the **correct sequence** to form complete fault chains*
  - *Tele-network ⟶ heterogeneous graph*
  - ***Link prediction** in fault chain*

Method	MRR	Hits@1	Hits@3	Hits@10
Random	58.2	56.2	56.2	62.5
MacBERT	65.9	62.5	65.6	68.8
TeleBERT	69.0	65.6	71.9	71.9
KTeleBERT-STL	73.6	71.9	71.9	78.1
w/o ANEnc	67.5	65.6	65.6	71.9
KTeleBERT-PMTL	87.3	84.4	87.5	93.8
KTeleBERT-IMTL	**94.8**	**93.8**	**93.8**	**100.0**

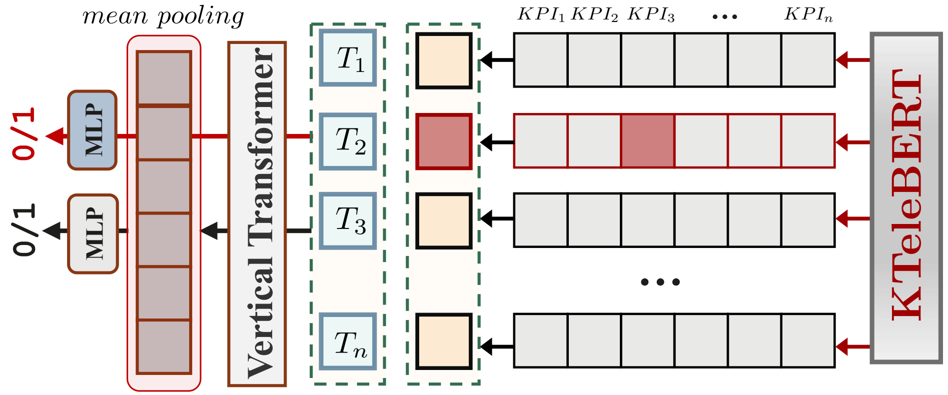


Task 4:

- ***Abnormal KPI Detection***
  - *Discover the abnormal value change within the **real-time KPI flow** of the tele-system*
  - *Multiple KPI indicators are recorded simultaneously on each NE (device)*
  - *Separate machine (**horizontal**) data points*
  - *continuous machine data (**vertical**) segments*



(a) KTeleBERT (w/ ANEnc).



# Abnormal Seg.	# Abnormal Poi.	# Normal Seg.	# Normal Poi.
4,510	7,512	7,864	90,912

Method	Recall Poi. (%)	Recall Seg. (%)
KTeleBERT (w/o ANEnc)	0.00	70.79
KTeleBERT (w/ ANEnc)	**56.57**	**91.48**

Future Work

- **Technical Routes :**
  - **GPT**✓V.S. **BERT**
  - Decoder-based models have shown greater potential in human-like question-answering and reasoning.
  - More training corpus and larger models
  - **TeleGPT**
- **Backbone :**
  - PanGu Chinese Language Model
- **Service Delivery Paradigm** :
  - Service embedding
  - Chain-of-thought
  - …

Thank you!

https://github.com/hackerchenzhuo/KTeleBERT

Tele-Knowledge Pre-training for Fault Analysis