



Tele-Knowledge Pre-training for Fault Analysis

Zhuo Chen^{1,4†}, Wen Zhang^{1,4†}, Yufeng Huang¹, Mingyang Chen¹, Yuxia Geng¹, Hongtao Yu¹, Zhen Bi¹, Yichi Zhang¹, Zhen Yao¹, Wenting Song², Xinliang Wu², Yi Yang², Mingyi Chen², Zhaoyang Lian², Yingying Li², Lei Cheng², Huajun Chen^{1,3,4‡}

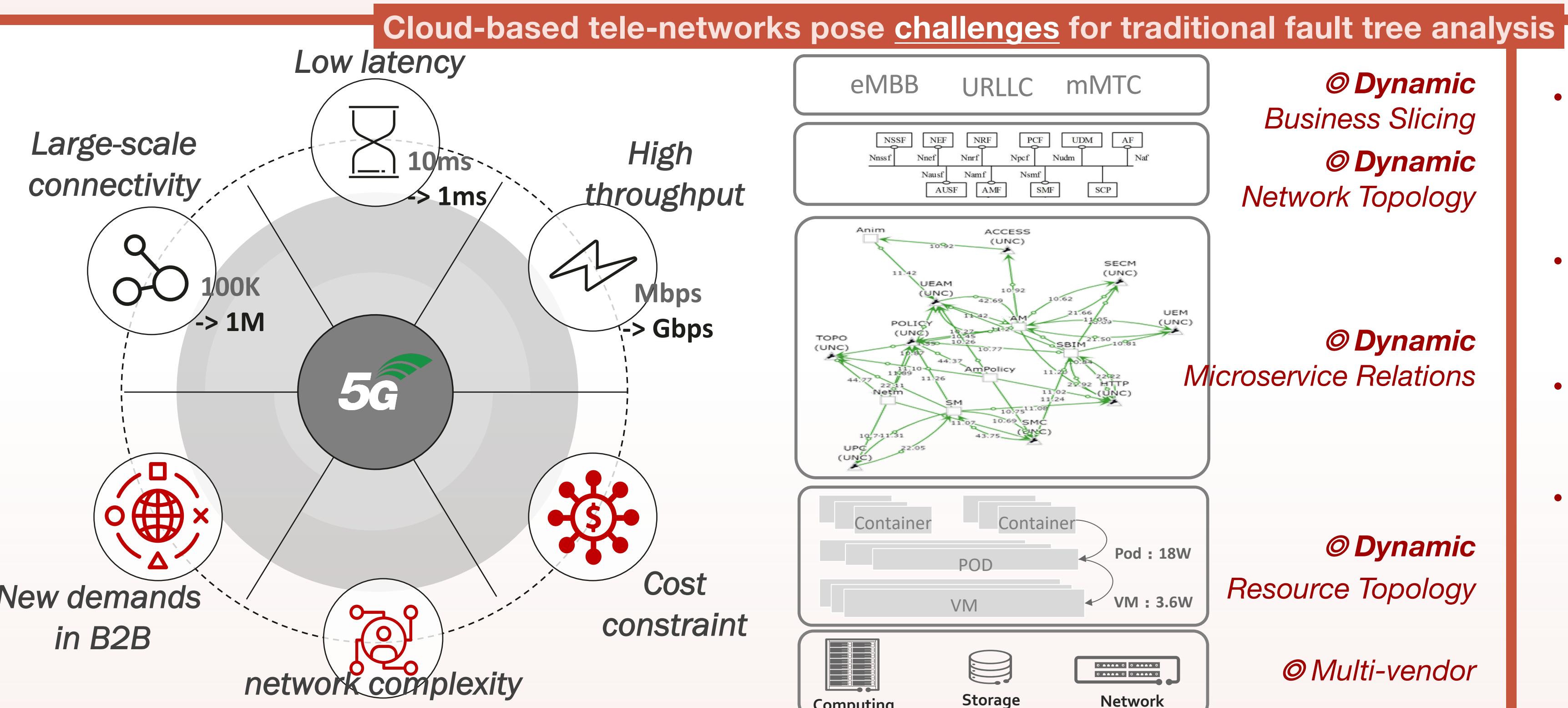
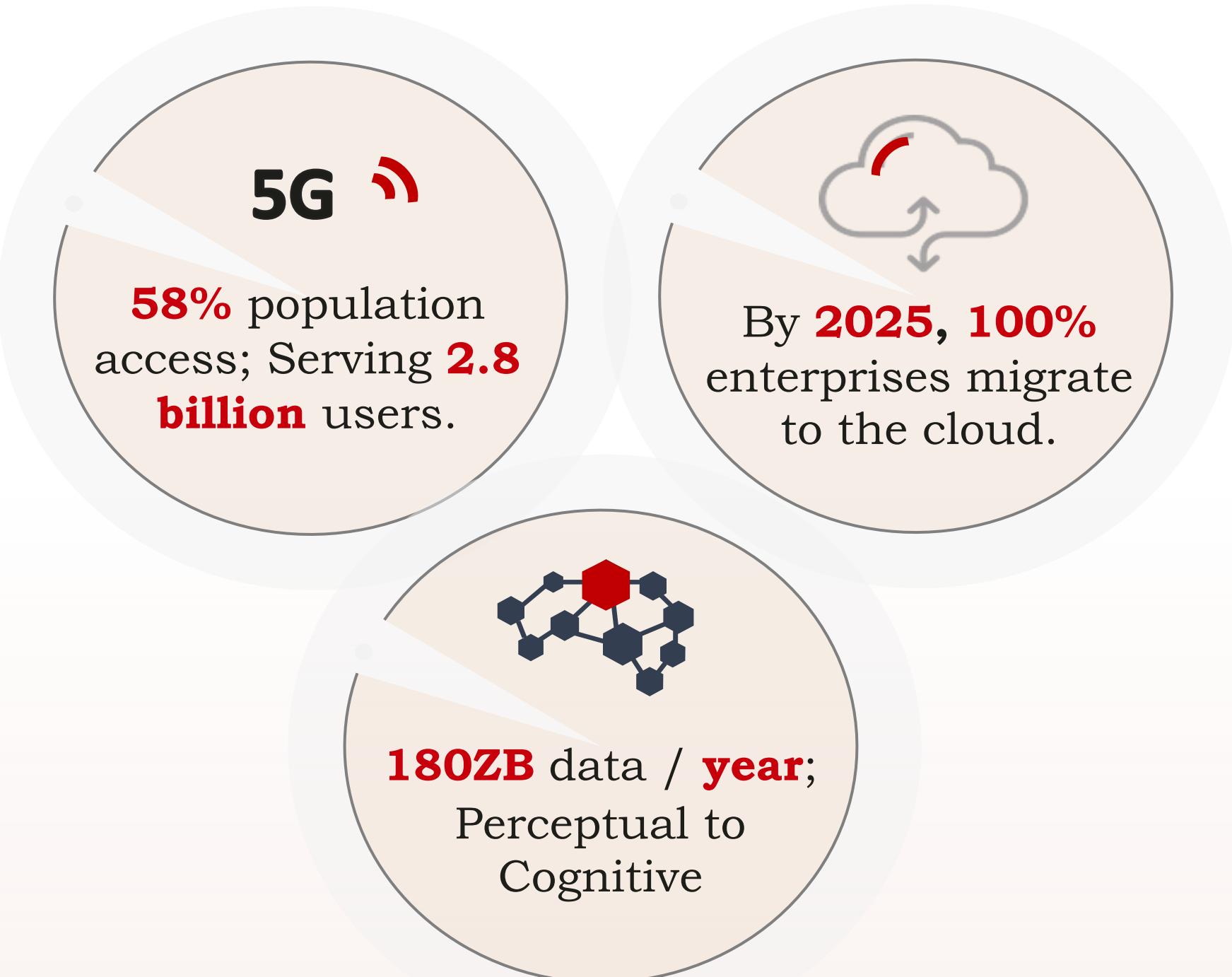
<https://github.com/hackerchenzhuo/KTeleBERT>

1 Zhejiang University 2 NAIE PDU, Huawei Technologies Co., Ltd

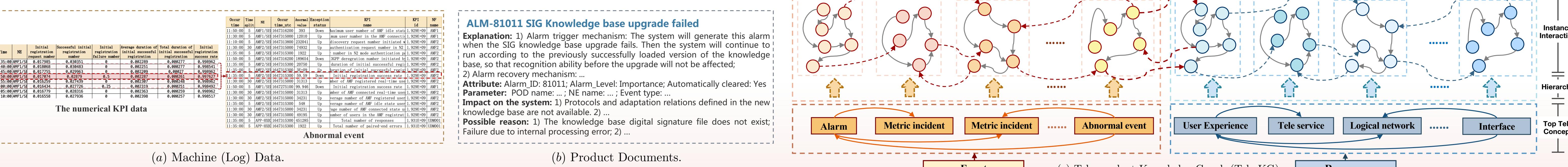
3 Donghai Laboratory 4 Alibaba-Zhejiang University Joint Institute of Frontier Technologies



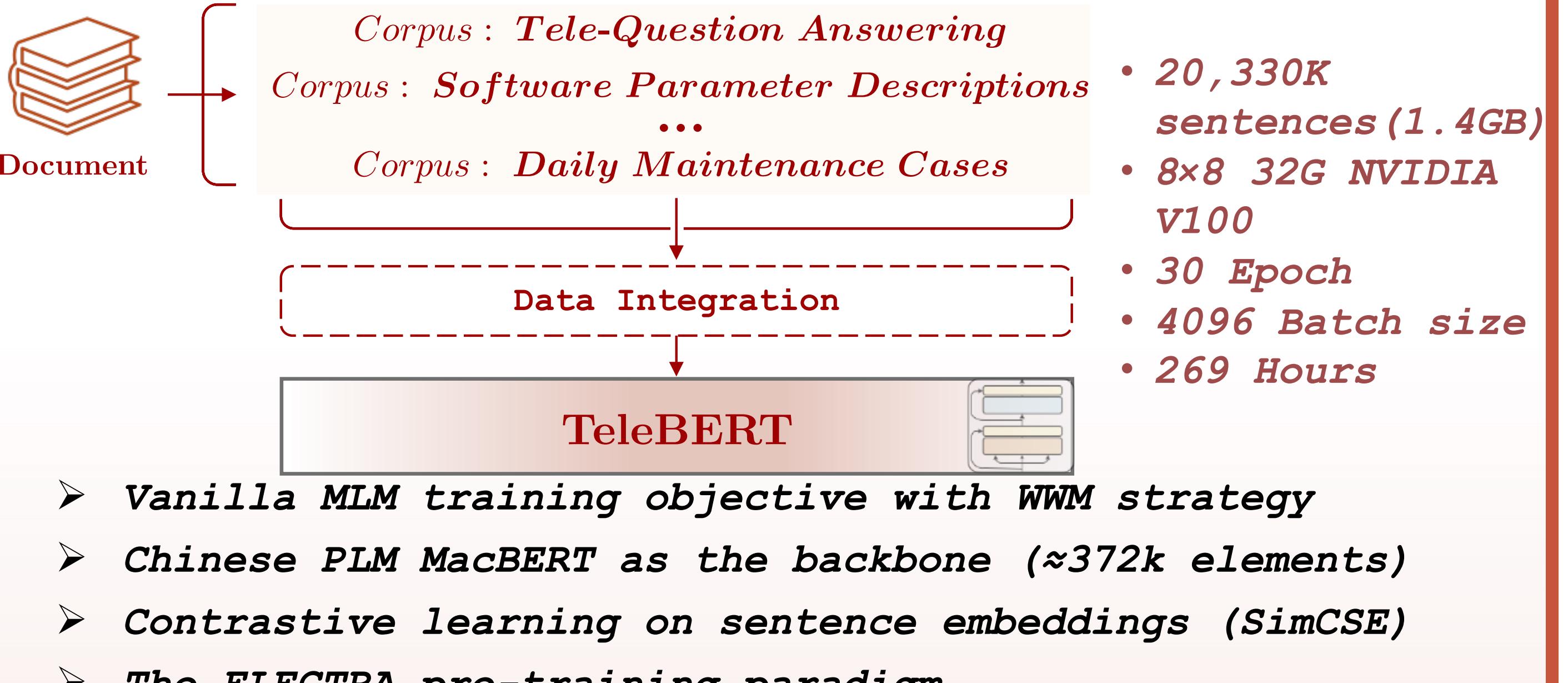
Background



- High Number of Alarms:** • 7,000+ alarms and 230+ alarm-related work orders per day
- Various fault types:** • 7 major categories and thousands of subcategories
- Difficult to Locate:** • Locating faults in network element links is difficult
- Difficult to Accumulate Experience:** • Rule-based Fault analysis tools are unable to handle new faults that have never occurred before

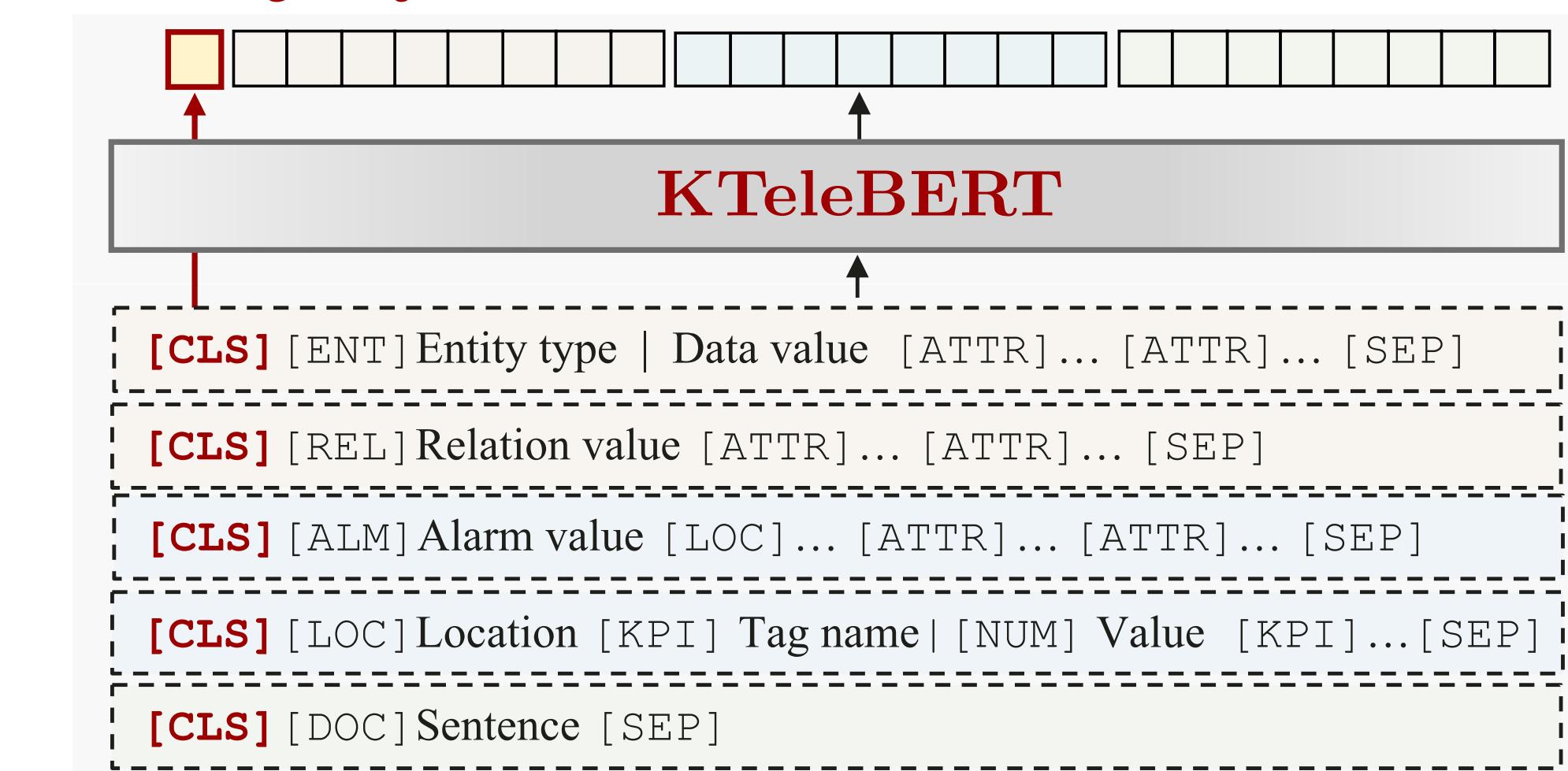


TeleBERT : Tele-Document Pre-training



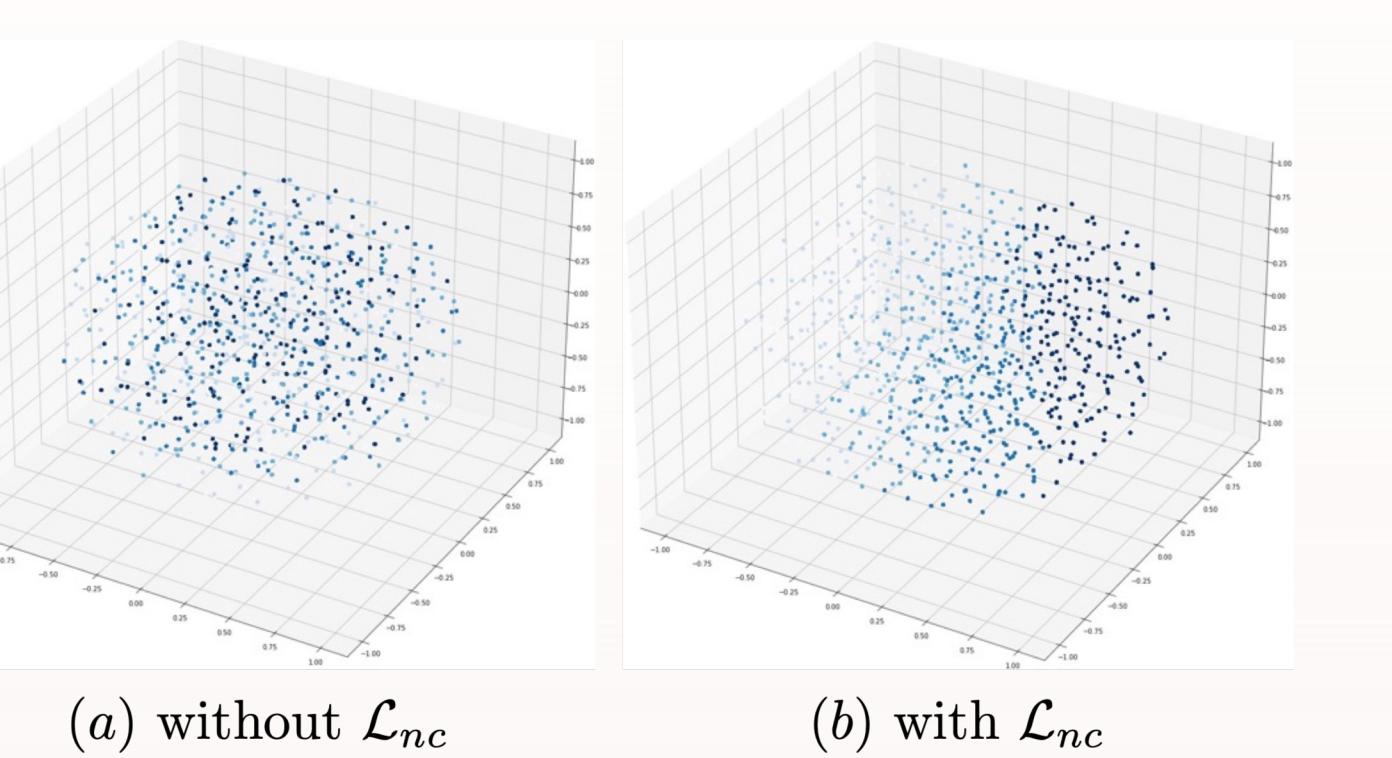
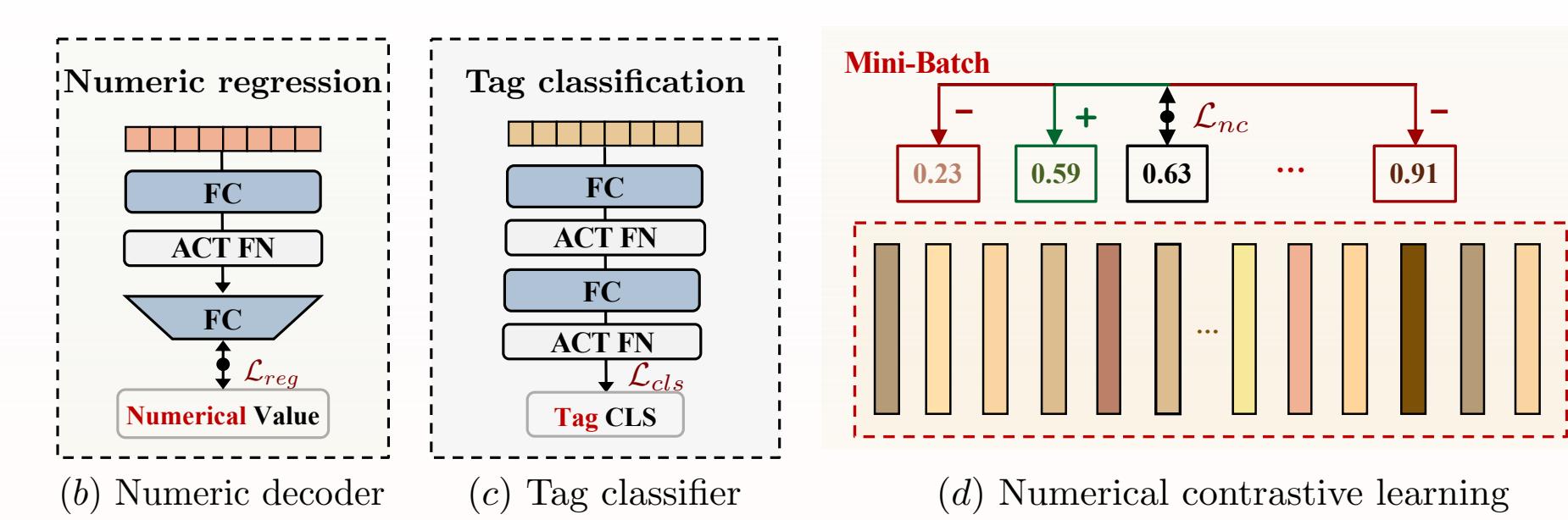
- Vanilla MLM training objective with WWM strategy
- Chinese PLM MacBERT as the backbone ($\approx 372k$ elements)
- Contrastive learning on sentence embeddings (SimCSE)
- The ELECTRA pre-training paradigm

• Unifying Modalities and Patterns → Implicit Knowledge Injection

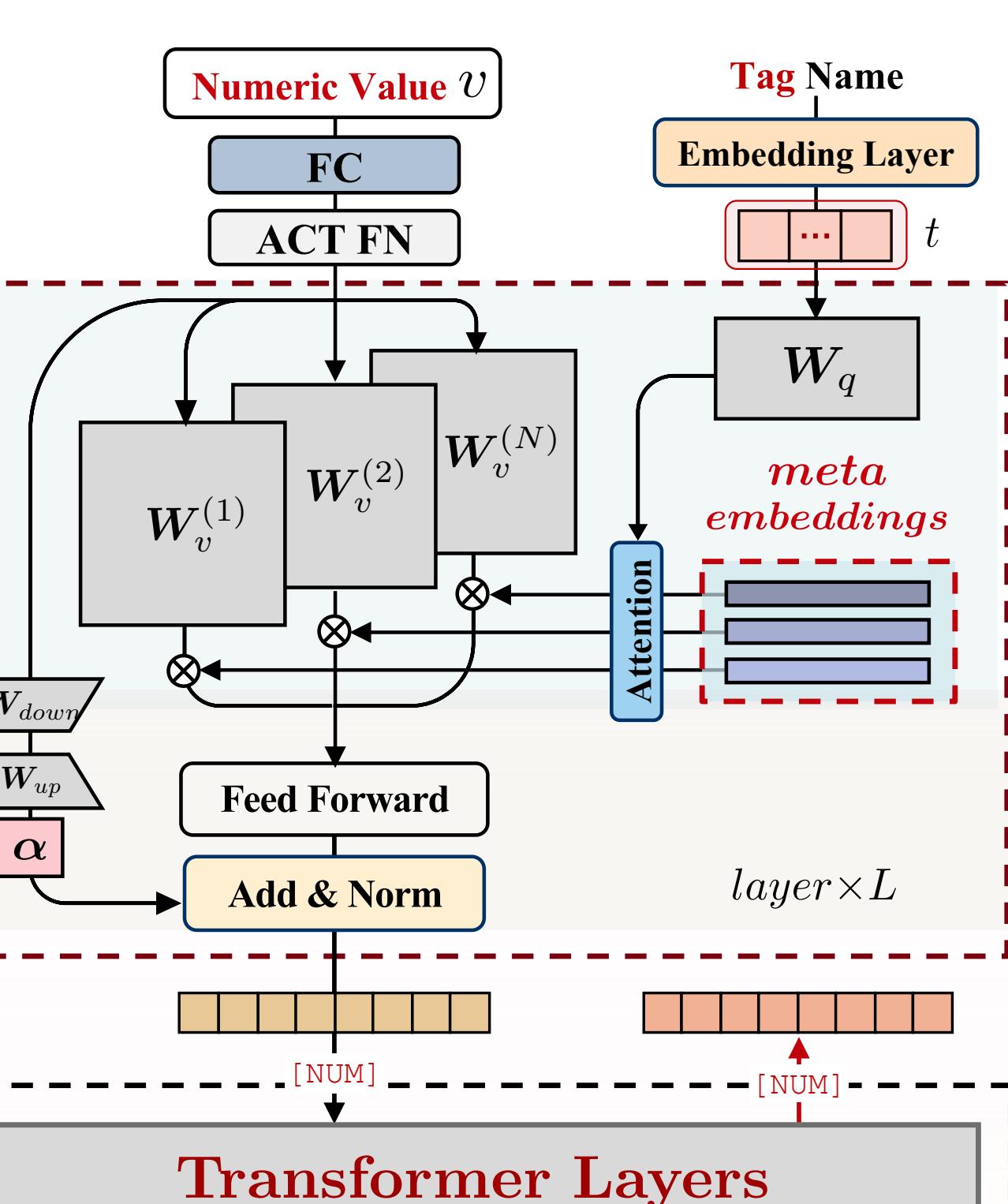


• Numerical Data Encoding → Machine (Log) Data

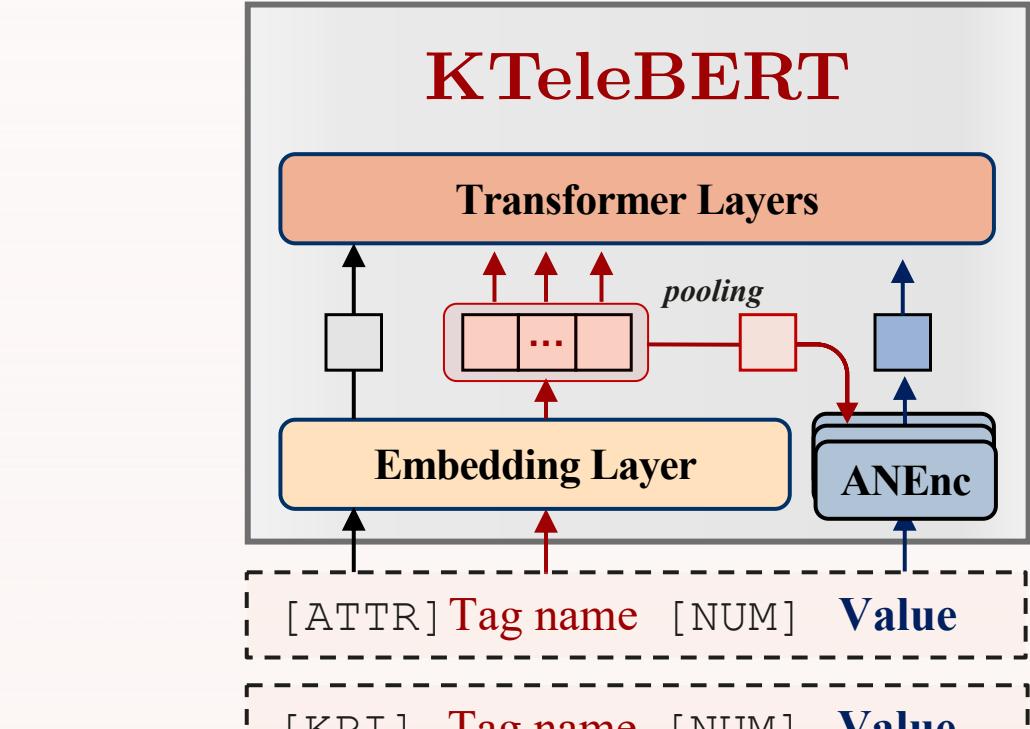
- ANEnc layer**
 - Attention-based numeric projection (ANP)
 - Learnable field-aware meta embeddings
 - Decouple domain knowledge
 - Corresponding Value conversion function
 - Numerical embedding transformation
 - Query projection
 - Tag name embedding conversion



- (a) Causal sentences extraction (200k)
 - Causal meaning words/phrases
 - Extraction rules (e.g., minimum length)
 - Tele-KG Serialization
- (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
 - Byte Pair Encoding (BPE)
 - domain-specific phrases or nouns
 - length: 2~4
 - Frequency > 8000
 - E.g., RAN, MML, MME, NF



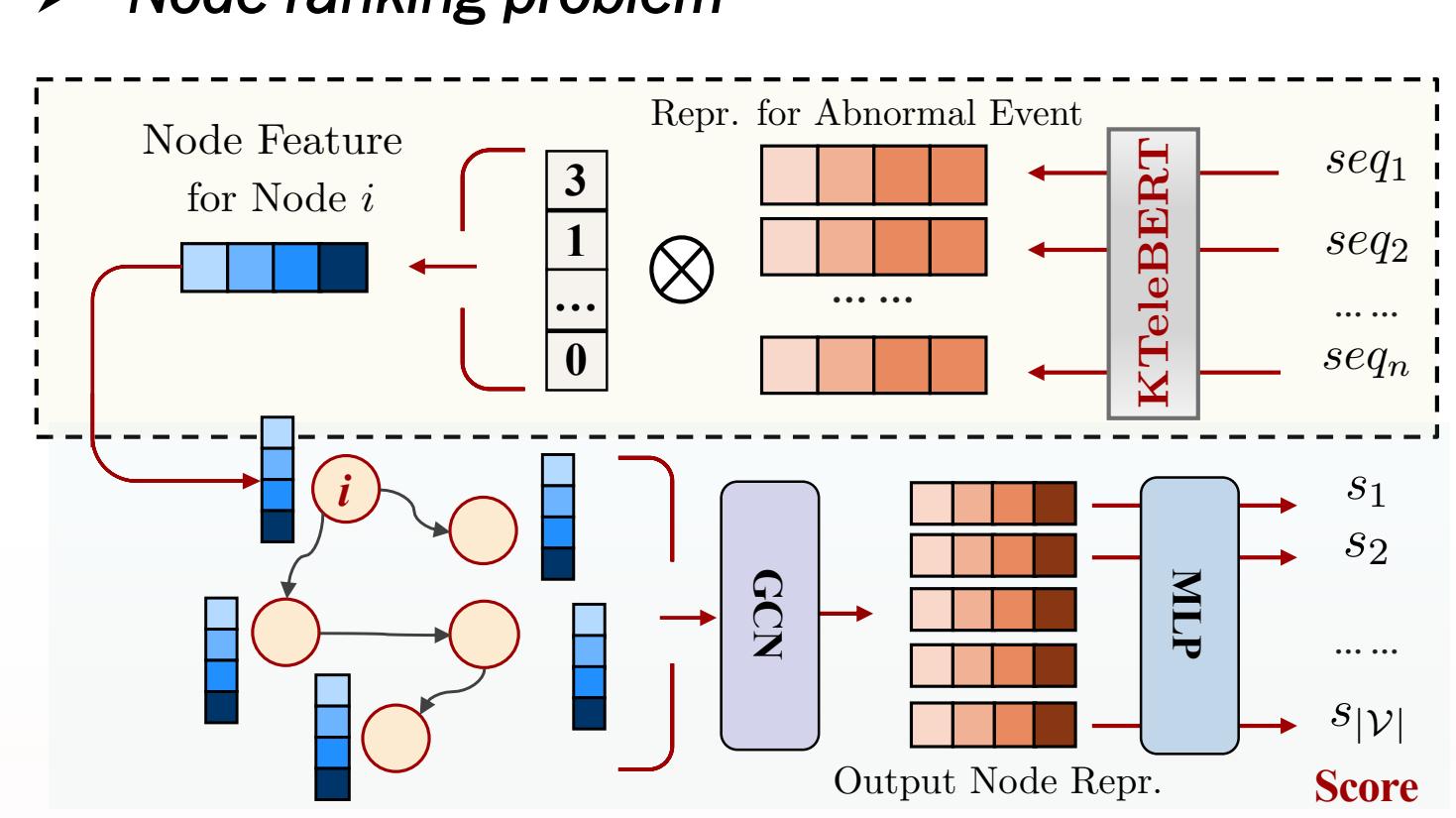
- Training Objective**
 - Numeric regression
 - Tag name classification
 - Numerical contrastive learning
 - Positive: closest value
 - Smooth numerical value changing
 - Stabilize the model



Fault Analysis Tasks

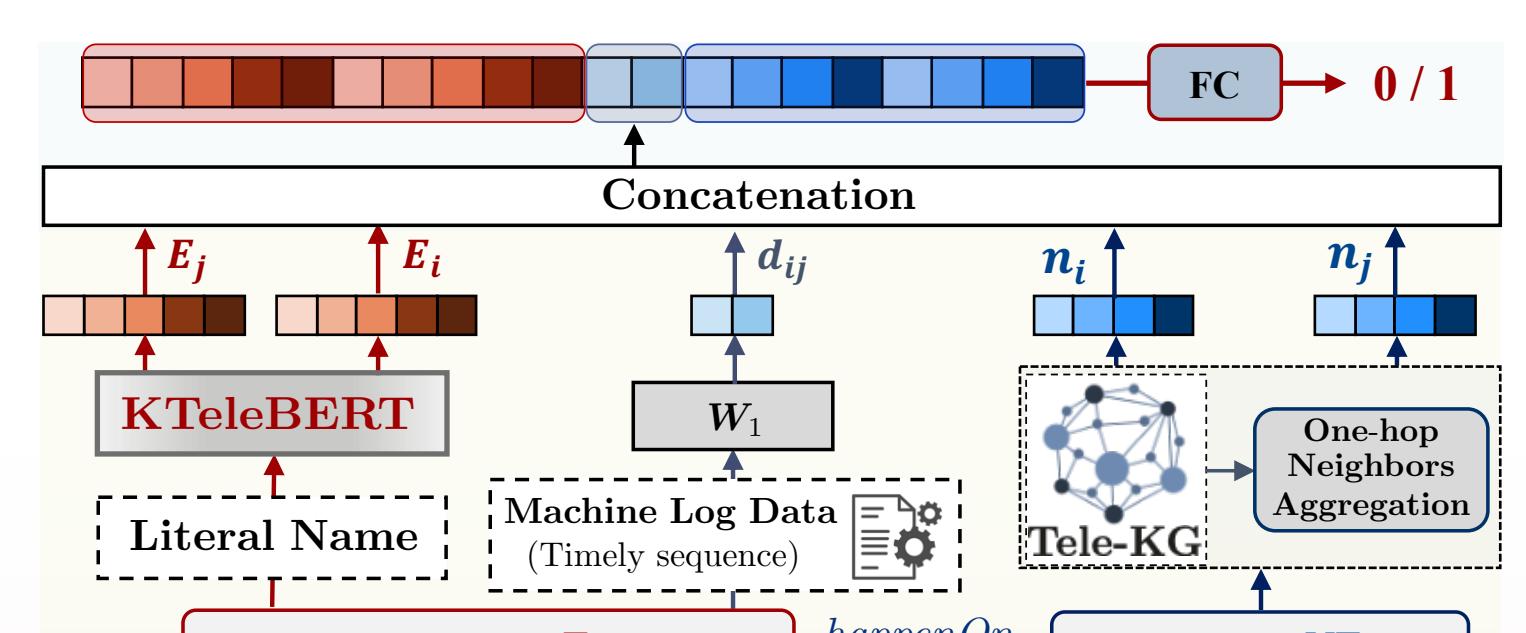
• Root-Cause Analysis

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



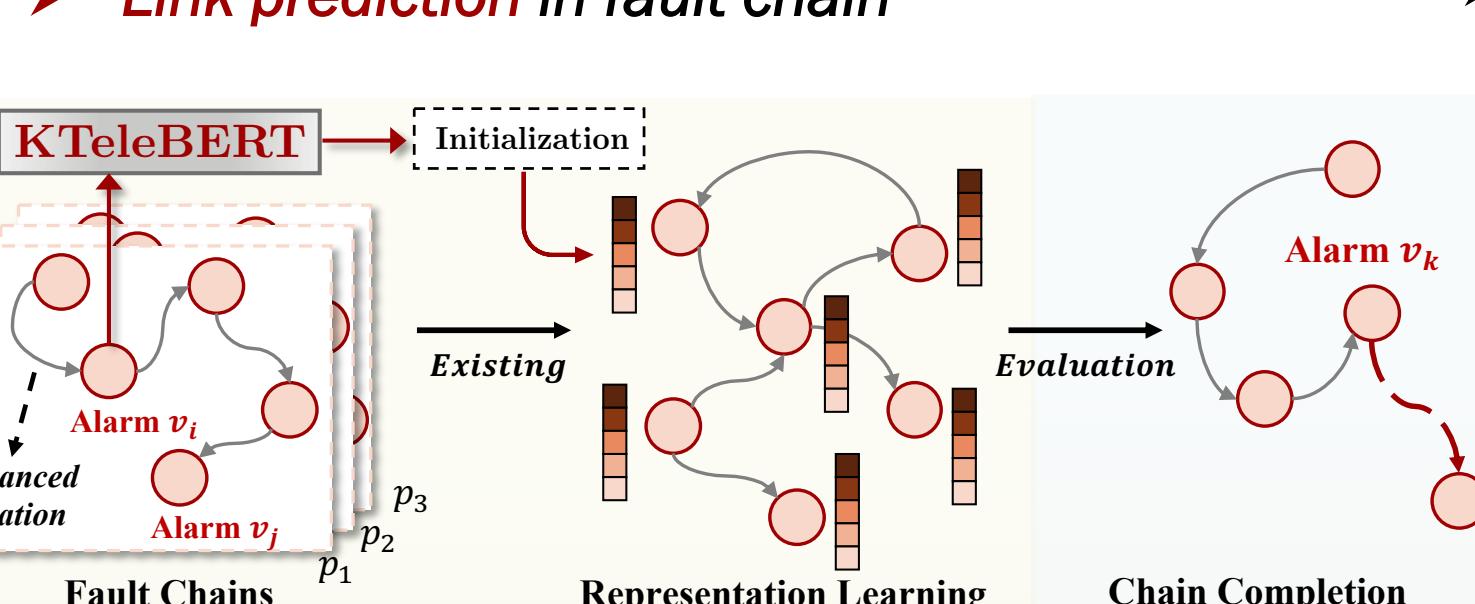
• Event Association Prediction

- Representing events in low-dimensional embeddings
- Learning the associations between events
- Trigger relation-specific space



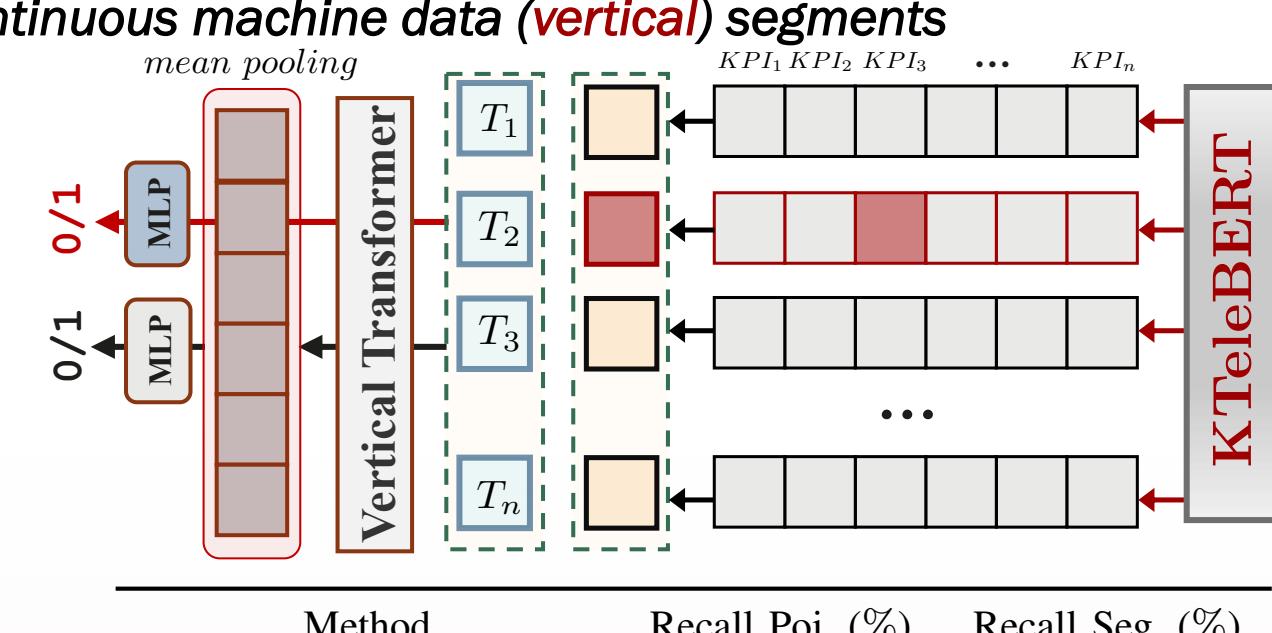
• Fault Chain Tracing

- Connect those faults in the correct sequence to form complete fault chains
- Tele-network → heterogeneous graph
- Link prediction in fault chain



• 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



Method	MRR	Hits@1	Hits@3	Hits@10
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 w/o ANEnc	2.06	63.66	83.21	91.87
KTeleBERT-PMTL	2.03	65.96	84.98	92.63
KTeleBERT-IMTL	2.02	64.78	85.65	91.13

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 w/o ANEnc	77.3	76.6	96.6	85.4
KTeleBERT-PMTL	68.5	68.8	99.1	81.3
KTeleBERT-IMTL	73.5	73.8	95.6	83.2

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 w/o ANEnc	73.6	71.9	71.9	78.1
KTeleBERT-PMTL	87.3	84.4	87.5	93.8
KTeleBERT-IMTL	94.8	93.8	93.8	100.0

