AIOps Challenge

2023 CCF国际AIOps挑战赛决赛
暨“大模型时代的AIOps”研讨会

KnowLog: Knowledge Enhanced Pre-trained Language Model for Log Understanding

徐波 东华大学副教授，复旦大学知识工场实验室副主任

ICSE 2024

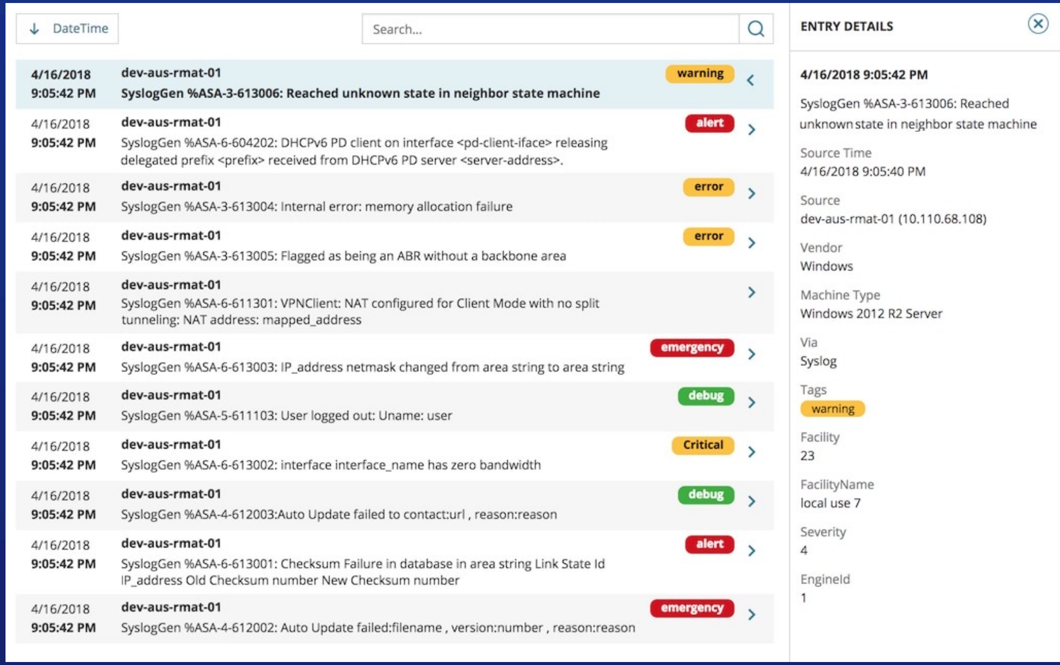
主办单位：中国计算机学会（CCF)、清华大学、中国建设银行股份有限公司、南开大学
承办单位：中国计算机学会互联网专委会、清华大学计算机科学与技术系、中国建设银行股份有限公司运营数据中心、南开大学软件学院、北京必示科技有限公司
赞助单位：华为技术有限公司、国网宁夏电力有限公司电力科学研究院、软通动力信息技术（集团）股份有限公司

目录
CONTENTS

研究背景

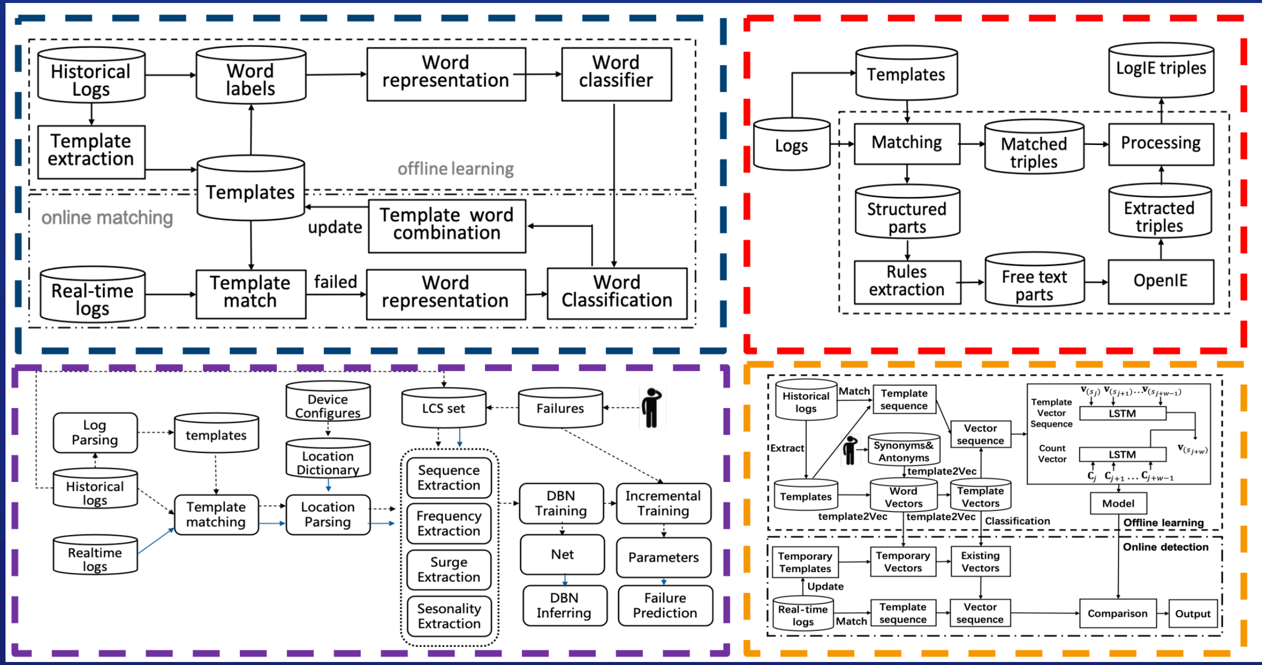
研究背景

- 在智能运维领域，随着系统的规模和复杂性不断增加，日志自动化分析的作用愈发重要
- 然而，每个日志分析任务都需要设计单独的模型，**缺乏统一**的处理框架[1]



运维团队使用日志文件监控系统运行状态

日志压缩框架 日志总结框架

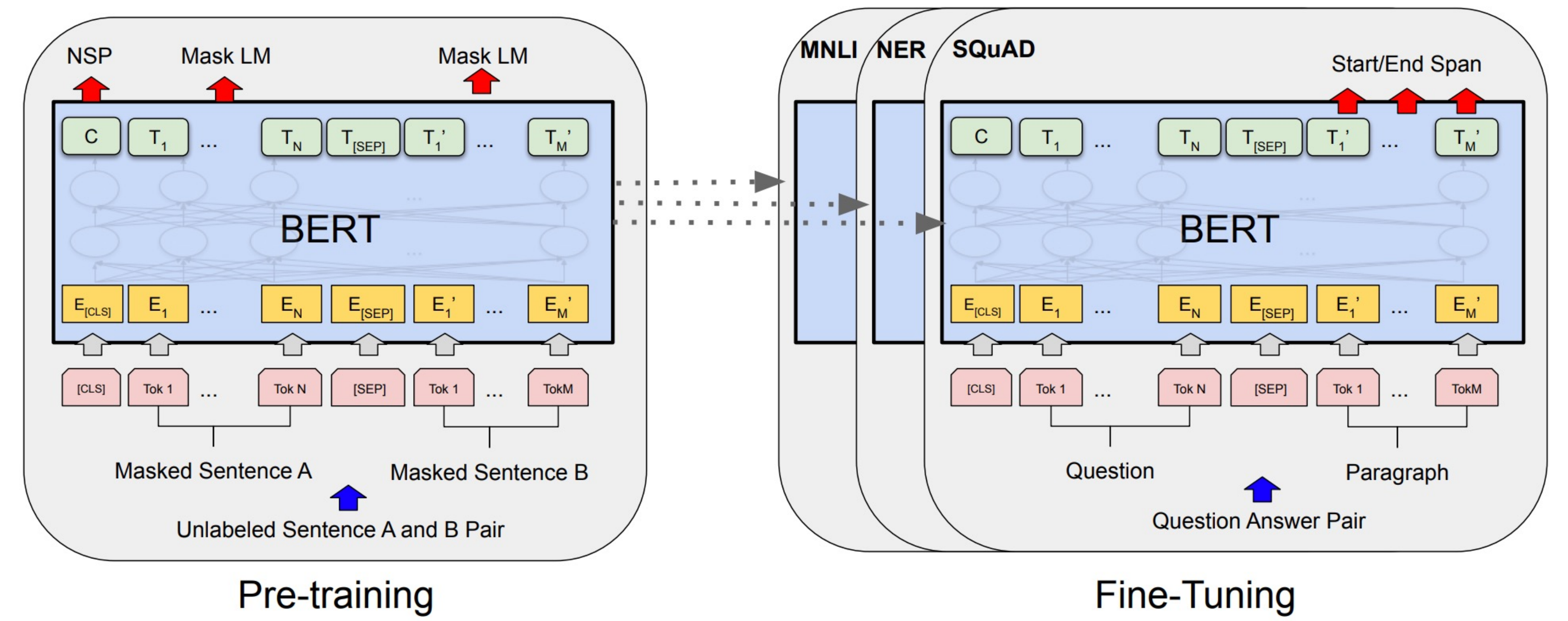


故障预测框架 异常检测框架

[1] Yichen Zhu et.al, UniLog: Deploy One Model and Specialize it for All Log Analysis Tasks. Arxiv 2021.12

研究背景

■ 在自然语言处理领域，以BERT[1]为代表的“**预训练+微调**”已经成为自然语言处理任务的统一处理框架



[1] Jacob Devlin et.al, BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Arxiv 2019.05

研究背景

AIOps Challenge

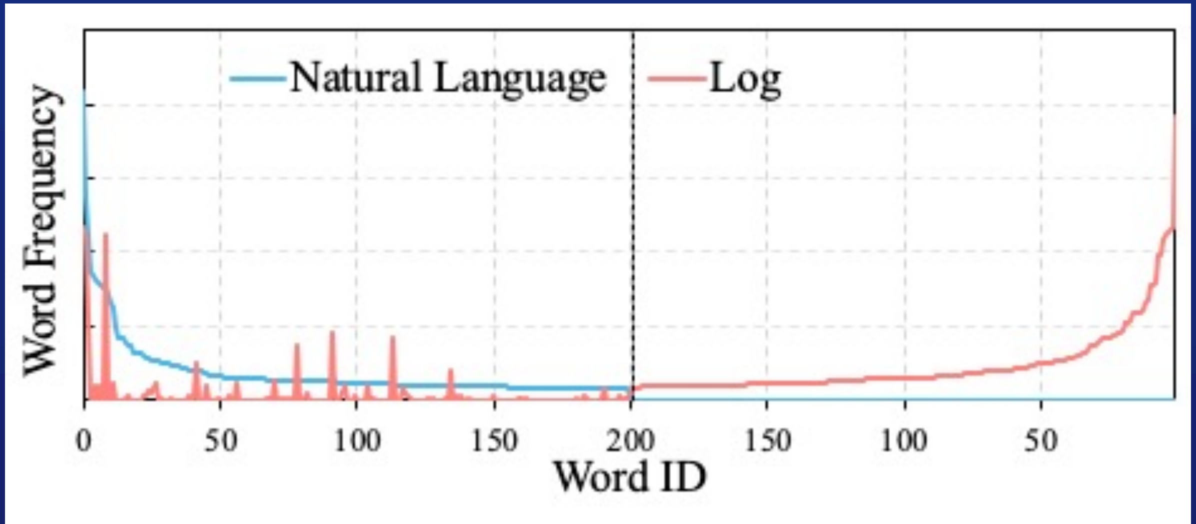
- 自然语言预训练模型**难以**表征日志
  - 日志是一种由模板和变量组成的半结构化语言（非自然语言）
  - 真实场景中日志的词汇分布同自然语言有着较大差异

日志

```
2023-01-14 23:05:14 INFO:  Reading data from /user/input/file.txt
2023-01-14 23:05:14 DEBUG: Setting block size to 1919810
2023-01-14 23:05:14 INFO:  Setting replication factor to 4
2023-01-14 23:05:14 ERROR: /user/input/file.txt does not exist
```

结构化数据

Date	Time	Level	Template	Parameters
2023/1/14	23:05:14	INFO	Reading data from <*>	['/user/input/file.txt']
2023/1/14	23:05:14	DEBUG	Setting block size to <*>	['1919810']
2023/1/14	23:05:14	INFO	Setting replication factor to <*>	['4']
2023/1/14	23:05:14	ERROR	<*> does not exist	['/user/input/file.txt']



挑战

AIOps Challenge

- 自然语言预训练模型在分析日志时存在以下问题：
  - 难以理解日志中特定术语的含义
    - 日志中包含大量特定术语，如缩略词，由于在自然语言中鲜有出现，这对于理解日志是一个挑战
  - 难以理解整条日志的含义
    - 日志通常精炼，缺少上下文信息难以充分理解完整语义信息
  - 难以理解不同厂商对同一日志的不同表达
    - 不同厂商或系统间的日志存在着语法差异

```
%OSPF-4-SYSLOG_SL_MSG_WARNING: OSPF-4-DUPRID: message repeated 1 times in last 16 sec
%OSPF-4-SYSLOG_SL_MSG_WARNING: OSPF-4-DUPRID: message repeated 1 times in last 19 sec
%OSPF-4-DUPRID:  ospf-1000 [8580] (default) Router 100.0.0.1 on interface Vlan1000 is using our routerid, packet dropped
%OSPF-4-DUPRID:  ospf-1000 [8991] (default) Router 100.0.0.4 on interface Vlan1000 is using our routerid, packet dropped
```

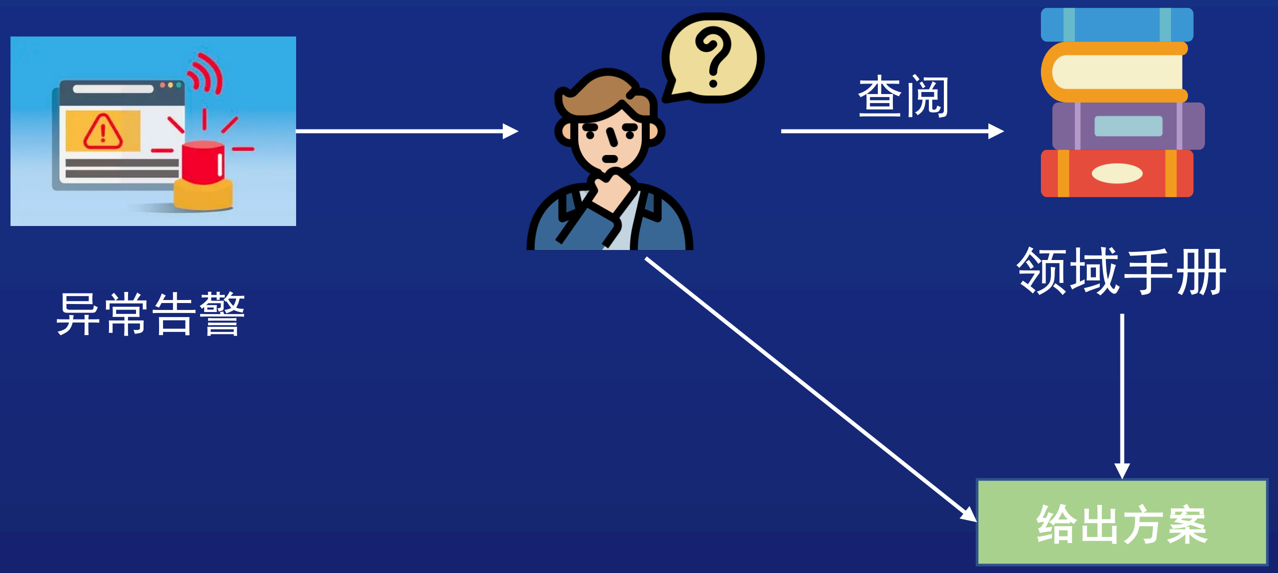
```
%%01INFO/4/SUPPRESS_LOG(l)[18]:Last message repeated 1 times.(InfoID=1077493797, ModuleName=SHELL, InfoAlias=LOGINFAILED)
%%01INFO/4/SUPPRESS_LOG(l)[7692]:Last message repeated 1 times.(InfoID=1077493797, ModuleName=SHELL, InfoAlias=LOGINFAILED)
%%01OSPF/4/CONFLICT_ROUTERID_INTF(l):CID=0x80820445;OSPF router ID conflict is detected on the interface.(ProcessId=1, RouterId=10.84.21.111,
AreaId=0.0.0.0, InterfaceName=10GE1/0/11, IpAddr=11.172.10.1, PacketSrcIp=11.172.10.2)
%%01OSPF/4/CONFLICT_ROUTERID_INTF(l):CID=0x80820445;OSPF router ID conflict is detected on the interface.(ProcessId=1, RouterId=10.84.21.111,
AreaId=0.0.0.0, InterfaceName=10GE1/0/11, IpAddr=11.172.10.1, PacketSrcIp=11.172.10.2)
```

知識工場

技术方案

技术方案：知识增强的日志预训练模型构建

- 受领域专家解决问题的思路启发：
  - 遇到不熟悉的日志通过查阅领域手册（外部知识）来解决问题

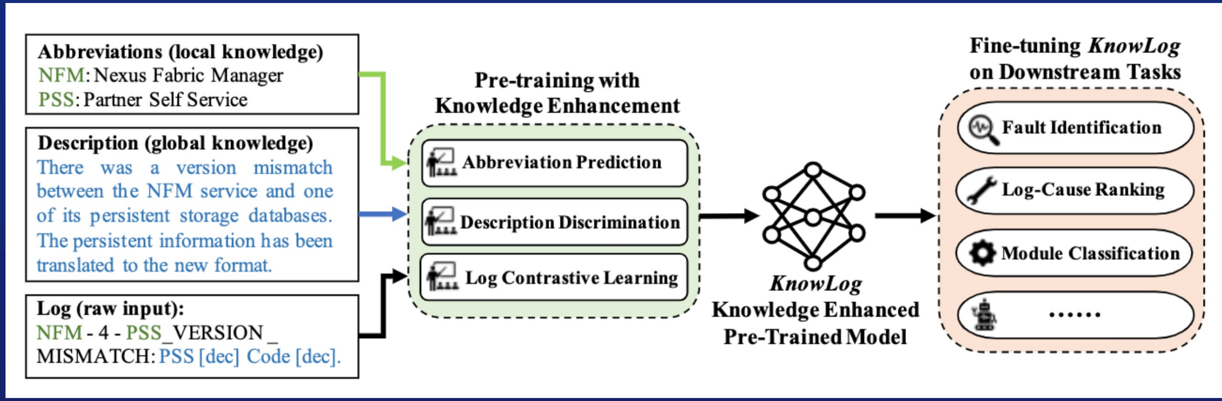


技术方案：知识增强的日志预训练模型构建

AIOps Challenge

■ 通过知识增强的手段为模型注入领域知识

- 知识来源：领域手册
- 局部知识：术语表
- 全局知识：日志模板描述



架构总览图

Module name representation	Module name expansion
AAA	Authentication, Authorization and Accounting
ACL	Access Control List
ANCP	Access Node Control Protocol
APMGR	Access Point Management
ARP	Address Resolution Protocol
ATK	ATK Detect and Defense
ATM	Asynchronous Transfer Mode
BFD	Bidirectional Forwarding Detection
BGP	Border Gateway Protocol

PIM/4/NBR_DOWN

Message

PIM/4/NBR_DOWN: In the VPN instance, a neighbor was deleted from the interface. (VPNName=[VPNName], NbrAd
LastHelloTime=[LastHelloTime]s)

Description

In the VPN instance, a neighbor was deleted from the interface.

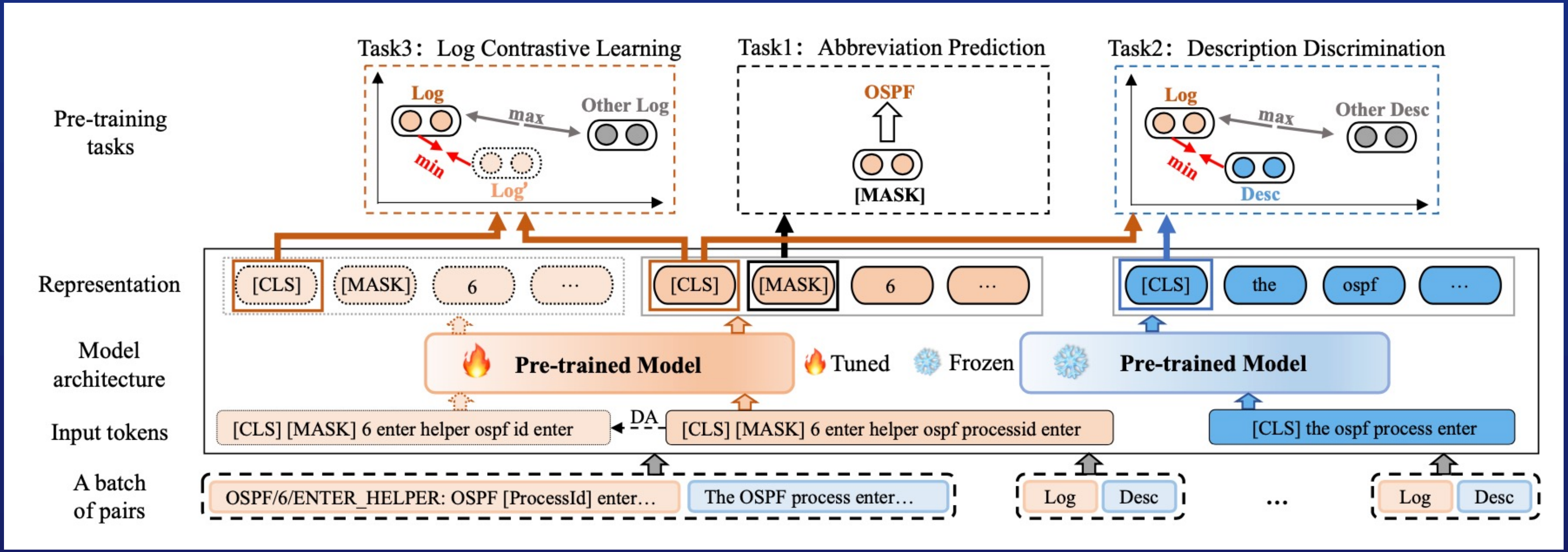
来自华为产品手册

技术方案：知识增强的日志预训练模型构建

AIOps Challenge

■ 通过自监督学习任务，强化预训练模型对于领域知识的习得

- 1）**局部知识**增强，设计**缩略词预测任务**，使模型能够理解缩略词。
- 2）**全局知识**增强，建立**日志和自然语言描述的对应关系**，拉近日志和描述的语义距离
- 3）**对比知识**学习：不同厂商间日志语法结构差异大，为了获得更通用的日志表征，对日志进行样本增强然后拉近日志间的语义距离



实验分析

实验分析

■ 数据集及下游任务

Table 1: Statistics of the dataset used for pre-training.

	Switches	Routers	Security	WLAN	All
Cisco	41,628	22,479	1,578	6,591	72,276
Huawei	6,418	4,980	3,737	1,001	16,136
H3C	2,171	2,364	1,852	1,261	7,648
All	50,217	29,823	7,167	8,853	96,060

预训练数据集

Table 2: Downstream tasks and their dataset statistics (Training/Validation/Testing Size).

		Switches	Routers	WLAN
Module Classification	Cisco	13,495/4,498/4,498	7,265/2,422/2,421	3,044/1,014/1,014
	Huawei	3,439/1,146/1,146	2,539/846/845	544/181/181
	H3C	1,241/413/413	1,336/445/444	724/241/241
Risk Log Identification	Huawei	788/263/262	502/167/166	379/126/125
Fault Phenomenon Identification	Huawei	362/120/120	-	-
Log and Description Semantic Matching	Cisco	49,954/16,651/16,651	26,975/8,992/8,991	7,910/2,636/2,636
	Huawei	7,702/2,567/2,567	5,977/1,992/1,991	1,202/400/400
	H3C	2,606/868/868	2,837/946/945	1,514/504/504
Log and Possible Cause Ranking	Huawei	3,851/1,283/1,283	3,097/1,032/1,032	602/200/200
Inter-vendor Module Matching	Huawei-Cisco	3,337/1,112/1,111	2,121/707/706	483/161/160
	Huawei-H3C	3,533/1,178/1,177	2,690/896/896	437/146/145
	Cisco-H3C	3,059/1,020/1,019	-	-

下游任务数据集

Table 15: Examples of Log-single tasks.

Tasks		Example
MC	**Input**	[MASK]/6/NOTIFY_RECV: The router received a NOTIFICATION from the peer. (Peer=[peer-address], SourceInterface=[SourceInterface], ErrorCode=[error-code], SubErrorCode=[sub-error-code], NotifyInfo=[notify-info], VpnInstance=[VpnInstance], ErrorData=[error-data])
	Output	BGP
	Input	[MASK]-3-DUPLICATE_IFINDEX:%s has %d duplicate ifIndices.
	Output	SNMP
RLI	**Input**	RRPP/2/MULMAST:OID [oid] A conflicting master node was detected on RRPP domain [domain-id] ring [ring-id].
	Output	True
	Input	RUMNG/4/RUPORTOPTPWRRESUME:OID [oid] Remote unit optical module recovered from power abnormal. (RemoteUnitEsn=[OCTET], InterfaceName=[OCTET], ReasonDescription=[OCTET])
	Output	False
FPI	**Input**	OSPF/4/CONFLICT_ROUTERID_INTF: OSPF router ID conflict is detected on the interface.(ProcessId=1, RouterId=10.26.09.101, AreaId=0.0.0.0, InterfaceName=10GE1/0/11, IpAddr=10.26.10.1, PacketSrcIp=10.26.10.2)
	Output	Router_id_conflict
	Input	IFNET/2/linkDown_activ: The interface status changes. (ifName=10GE1/0/11, AdminStatus=DOWN, OperStatus=DOWN, Reason=The interface is shut down, mainIfname=10GE1/0/11)
	Output	Trunk_link_down & Physical_link_down

Table 16: Examples of Log-pair tasks.

Tasks		Example
LDSM	**Input**	[ARP/4/ARP_VLAN_SPEED_LMT: The VLAN's ARP packet speed exceeded the configured speed limit value. (SuppressValue=[SpeedLmtValue], Vlan=[VlanId]) , The transmit rate of ARP packets in a VLAN exceeded the configured rate limit in the VLAN.]
	Output	True
	Input	[(ARP/4/ARP_VLAN_SPEED_LMT: The VLAN's ARP packet speed exceeded the configured speed limit value. (SuppressValue=[SpeedLmtValue], Vlan=[VlanId]) , A received ARP packet was not an ARP reply packet in response to the ARP request packet sent by the device.]
	Output	Fasle
LPCR	**Input**	BGP/3/FSM_UNEXPECT: FSM received an unexpected event. (FSM=[fsm-name], PreState=[prev-state], CurrState=[curr-state], InputEvent=[input])
	Output	It is caused by an internal error of the system.
	Input	BGP/2/hwBgpPeerSessionExceed_clear: The number of BGP peer sessions decreased below the maximum number. (MaximumNumber=[MaximumNumber], CurrentNumber=[CurrentNumber])
	Output	The number of BGP peer sessions fell below the upper limit.
IVMM	**Input**	[[MASK]/2/hwBgpPeerSessionExceed_active: The number of BGP peer sessions exceeded the maximum number. (MaximumNumber=[MaximumNumber]) , [MASK]-3-MAXPFXEXCEED:Number of prefixes received from %s%s%s afi %d: %d exceeds limit %d)]
	Output	True
	Input	[[MASK]-3-MAXPFXEXCEED:Number of prefixes received from %s%s%s afi %d: %d exceeds limit %d) , [MASK]/3/hwTelnetLoginFailed_clear: The telnet user login-failed alarm was cleared.]
	Output	False

下游任务举例

实验分析

AIOps Challenge

■ 知识增强的日志预训练模型**显著优于**BERT等通用预训练语言模型

Table 3: Results on Module Classification and Risk Log Identification.

Methods	MC (Accuracy/Weighted F1)									RLI (Precision/Recall/F1)		
	Cisco			Huawei			H3C			Huawei		
	Switches	Routers	WLAN	Switches	Routers	WLAN	Switches	Routers	WLAN	Switches	Routers	WLAN
CNN	56.89/56.85	57.46/54.92	53.55/51.89	74.52/73.95	72.78/72.23	73.48/71.47	69.49/67.55	70.72/69.71	74.27/72.87	0.63/0.62/0.63	0.62/0.59/0.61	0.68/0.69/0.68
BiLSTM (Attention)	55.74/55.63	57.17/56.76	53.25/52.38	76.52/75.49	73.96/73.30	73.48/72.51	70.21/68.45	71.40/69.93	74.69/73.56	0.68/0.64/0.66	0.57/0.62/0.59	0.67/0.72/0/69
UniLog	63.83/63.45	64.60/63.44	62.13/61.18	83.07/82.11	81.30/79.57	88.95/87.68	81.60/79.80	79.28/77.75	80.50/78.92	0.70/0.67/0.69	0.76/0.72/0.74	0.84/0.80/0.82
BERT	62.67/61.38	62.72/62.60	61.63/60.87	82.37/81.20	81.18/79.20	86.19/84.89	81.11/79.78	77.93/76.05	80.08/76.90	0.69/0.61/0.64	0.70/0.73/0.71	0.80/0.82/0.81
KnowLog (BERT)	**68.03/67.80**	**70.05/69.22**	**66.57/66.29**	**86.13/85.36**	**86.39/85.23**	**88.95/88.49**	**82.57/80.95**	**80.86/79.19**	**81.33/79.18**	**0.75/0.70/0.73**	**0.76/0.78/0.77**	**0.85/0.89/0.87**
RoBERTa	62.72/62.58	63.90/63.08	60.95/60.77	81.50/80.64	81.18/79.20	86.19/85.01	81.35/79.73	78.60/76.81	80.08/77.97	0.70/0.64/0.67	0.70/0.73/0.71	0.80/0.82/0.81
KnowLog (RoBERTa)	**68.32/67.88**	**71.62/70.89**	**67.16/66.72**	**86.39/85.30**	**86.27/84.93**	**88.95/87.68**	**82.08/80.45**	**80.63/79.33**	**80.50/78.17**	**0.81/0.72/0.77**	**0.79/0.82/0.81**	**0.81/0.87/0.84**

Table 4: Results on Log and Description Semantic Matching and Log and Possible Cause Ranking.

Methods	LDSM (Accuracy/Weighted F1)									LPCR (Precision@1/MRR)		
	Cisco			Huawei						Huawei		
	Switches	Routers	WLAN	Switches	Routers	WLAN	Switches	Routers	WLAN	Switches	Routers	WLAN
CNN	84.04/84.04	80.99/80.99	72.15/72.16	86.05/86.05	82.37/82.30	72.75/72.75	83.29/83.19	83.60/83.59	82.54/82.42	-	-	-
BiLSTM (Attention)	89.45/89.44	85.42/85.41	76.86/76.86	87.85/87.85	84.43/84.40	72.75/72.73	80.88/80.83	83.81/83.80	82.74/82.72	-	-	-
UniLog	93.90/93.90	92.07/92.07	83.99/84.00	95.01/95.01	93.17/93.17	86.25/86.15	93.32/93.32	92.38/92.38	90.87/90.85	0.894/0.934	0.899/0.939	0.875/0.923
BERT	93.06/93.06	90.01/90.00	79.74/79.74	93.18/93.18	90.06/90.05	79.75/79.74	87.44/87.41	88.25/88.25	83.93/83.81	0.884/0.928	0.876/0.923	0.826/0.891
KnowLog (BERT)	**98.02/98.02**	**97.56/97.56**	**93.51/93.51**	**97.20/97.20**	**96.74/96.74**	**93.50/93.49**	**95.97/95.97**	**96.30/96.30**	**93.45/93.44**	**0.952/0.972**	**0.946/0.968**	**0.841/0.897**
RoBERTa	93.03/93.03	89.26/89.24	78.11/78.10	92.82/92.83	90.31/90.31	80.50/80.50	87.44/87.42	89.31/89.31	83.13/83.10	0.895/0.938	0.862/0.906	0.841/0.894
KnowLog (RoBERTa)	**96.56/96.56**	**96.32/96.32**	**93.25/93.25**	**97.20/97.20**	**96.23/96.23**	**93.25/93.24**	**95.05/95.04**	**96.08/96.08**	**94.84/94.84**	**0.935/0.962**	**0.935/0.963**	**0.861/0.910**

知識工場

实验分析

- 知识增强的日志预训练模型在**低资源场景**下具有显著优势

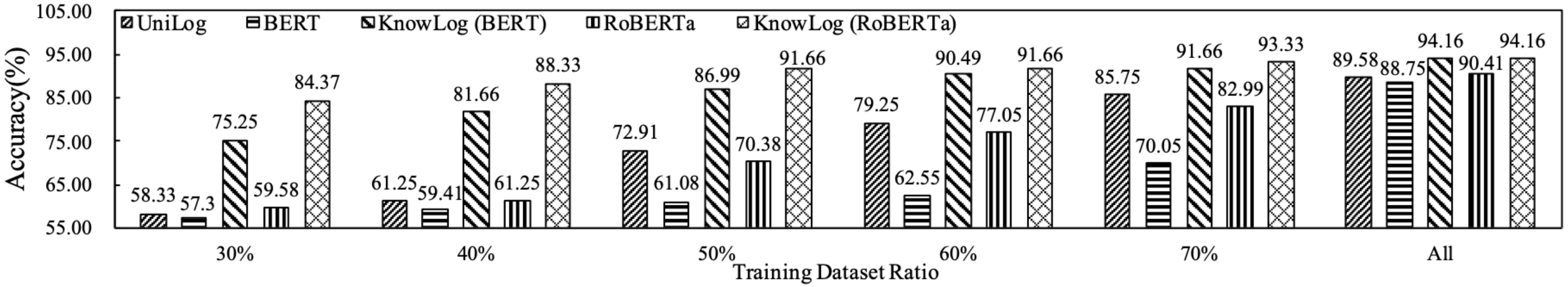


Figure 3: Results of KnowLog under different ratios of training dataset on the FPI task.

实验分析

AIOps Challenge

- 知识增强的日志预训练模型在**跨厂商迁移场景**下具有显著优势
- 将**领域术语**加入预训练模型词表中能够提升模型效果

Table 6: Results of transfer learning experiments on the task of LDSM. Left side of → indicates source dataset for training and right side indicates target dataset for testing.

Methods	Huawei → Cisco			Cisco → Huawei		
	Switches	Routers	WLAN	Switches	Routers	WLAN
CNN	62.25/61.59	62.05/61.55	52.35/52.07	69.61/69.40	62.13/62.13	60.25/60.09
BiLSTM (Attenion)	64.00/63.96	60.94/60.85	55.58/55.40	72.03/71.85	66.60/66.52	59.25/59.25
UniLog	77.57/77.40	77.33/77.27	66.69/66.70	90.49/90.49	87.74/87.74	83.75/83.73
BERT	71.46/71.13	72.97/72.88	61.38/61.36	83.73/83.72	84.53/84.52	72.75/72.25
KnowLog (BERT)	**86.63/86.59**	**87.22/87.19**	**80.39/80.04**	**95.56/95.56**	**93.17/93.16**	**96.25/96.25**
RoBERTa	71.95/71.74	73.57/73.41	61.53/61.54	84.03/84.02	84.63/84.62	73.25/73.25
KnowLog (RoBERTa)	**84.20/84.04**	**86.52/86.43**	**80.35/80.00**	**96.34/96.34**	**93.62/93.62**	**96.75/96.74**

Table 7: Results of whether abbreviations join the vocabulary.

Tasks		Abbr Not in Vocab	Abbr In Vocab
LDSM (Huawei)	Switches	96.61/96.61	**97.20/97.20**
	Routers	95.28/95.28	**96.74/96.74**
	WLAN	90.50/90.50	**93.50/93.49**
IVMM (Huawei-Cisco)	Switches	78.94/78.92	**79.30/79.28**
	Routers	75.50/75.51	**78.61/78.61**
	WLAN	86.88/86.88	**87.50/87.51**

知識工場

实验分析

AIOps Challenge

- 经过知识增强后，日志和其对应的自然语言描述的表征更为接近

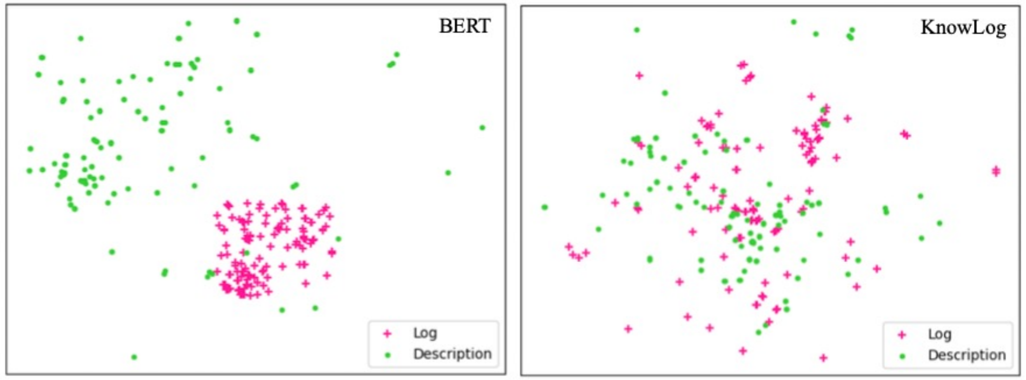


Figure 5: The representation visualization of logs and the corresponding descriptions.

Table 9: Qualitative examples of KnowLog and baselines. The input is a Log-NL or Log-Log pair and the Score indicates the cosine similarity.

Label	Examples	Models	Score
Match	**Log**: BGP/4/ASPATH_EXCEED_MAXNUM: The number of AS-PATHs exceeded the limit([limit-value]). (Operation=[STRING]) **NL**: The number of AS-Paths exceeded the maximum value.	BERT UniLog KnowLog	0.7250 0.7061 0.8006
UnMatch	**Log**: BGP/4/ASPATH_EXCEED_MAXNUM: The number of AS-PATHs exceeded the limit([limit-value]). (Operation=[STRING]) **NL**: The OSPF process successfully exited from GR.	BERT UniLog KnowLog	0.5715 0.3008 0.0056
Match	**Log1**: DEVM/3/hwRemoteFaultAlarm_active(l): The remote fault alarm has occurred. (IfIndex=27, IfName=10GE1/0/17) **Log2**: DEVM/3/hwRemoteFaultAlarm_active: The remote fault alarm has occurred.	BERT UniLog KnowLog	0.9550 0.8031 0.9750
UnMatch	**Log1**: BGP/4/ASPATH_EXCEED_MAXNUM: The number of AS-PATHs exceeded the limit([limit-value]). (Operation=[STRING]) **Log2**: DEVM/3/hwRemoteFaultAlarm_active: The remote fault alarm has occurred.	BERT UniLog KnowLog	0.8338 0.2997 0.1514

论文总结

总结

AIOps Challenge

- 提出了KnowLog，一个基于领域知识增强的日志预训练模型，可以更好的表征日志用在不同的自动化日志分析任务中；
- 通过设计合理有效的预训练任务使得模型融入领域知识，相较于通用预训练语言模型，KnowLog具备更好的日志理解能力；
- 融入领域知识的预训练模型在低资源和跨厂商迁移的场景中有显著提升。

知識工場

AIOps Challenge

2023 CCF国际AIOps挑战赛决赛暨“大模型时代的AIOps”研讨会

THANKS