



# Network anomaly detection based on logs

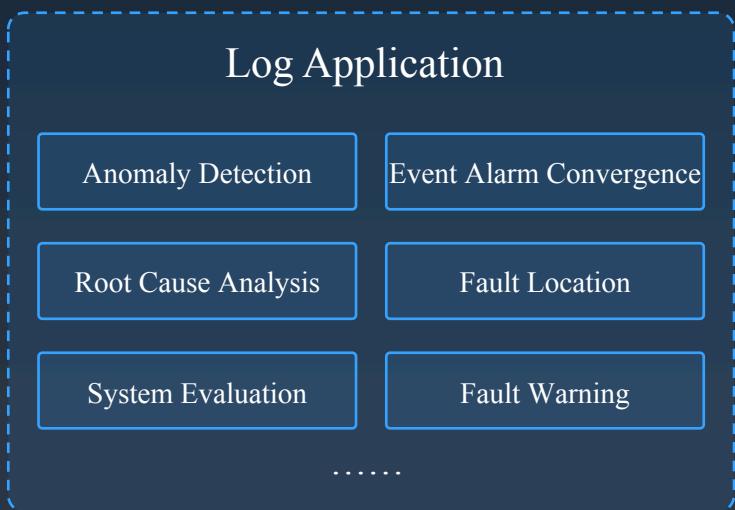
---



Chin : Qin YuKun , Song Yong , Xie Yuchen , Wei QiangShen

# Logs are valuable for network maintenance with existing challenges

Huge chunks of data was generated by Network Equipment and these information can be used to detect existing faults to improve the network maintenance.

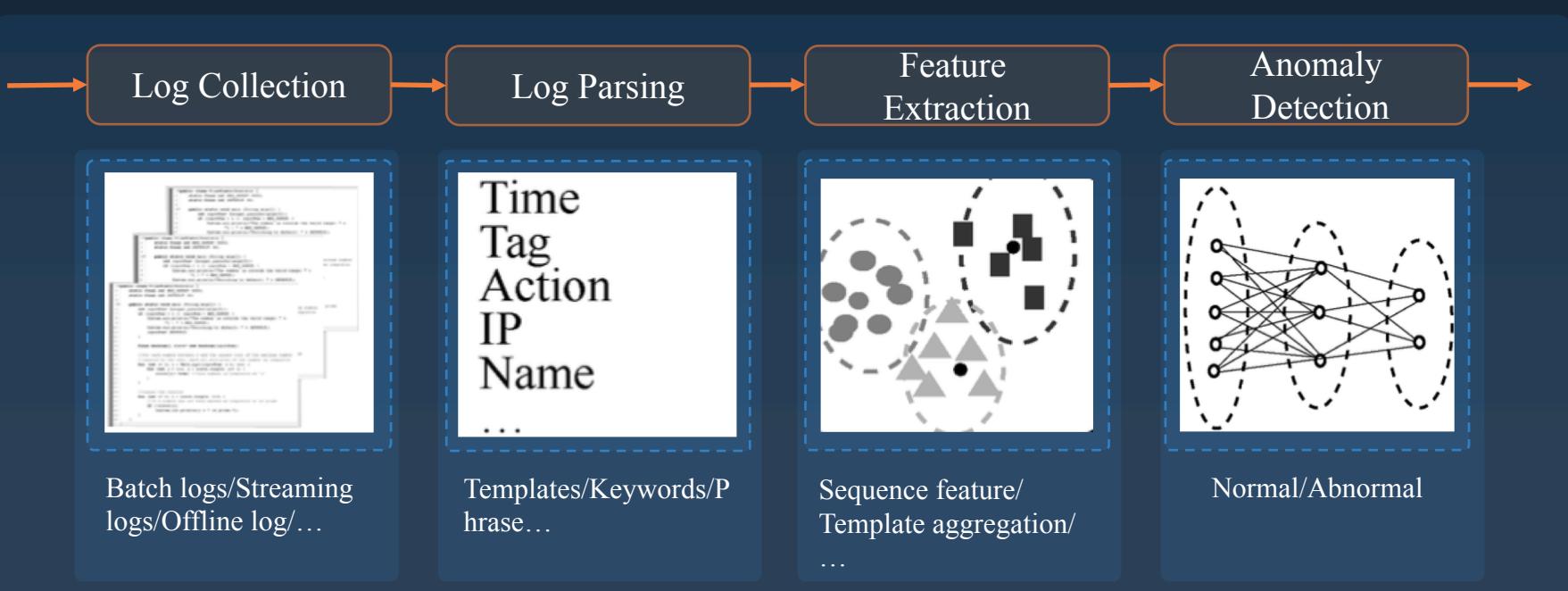


## Challenges

- Increasingly vast amount ( 1GB/H+ )
- Various types (Exchanger、 Server、 Router, etc. )
- Unstructured data
- Unsupervised / No-Label
- Tradeoff between effect and efficiency

# Ensemble Solution: NLP & Abnormal Detection

Network Anomaly Detection based on Logs was composed by NLP and abnormal detection.



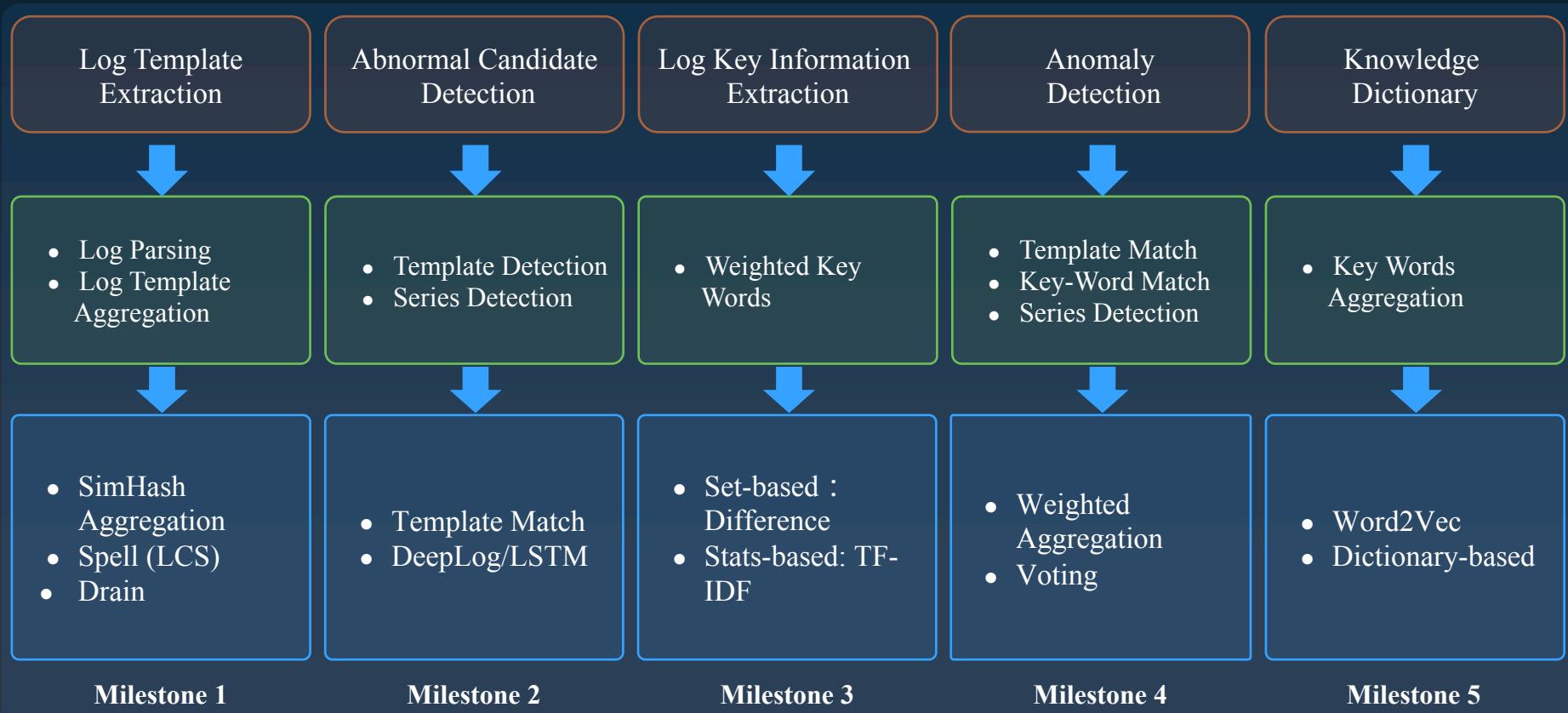
# Key-word Dictionary Construction is the most important part

Although there are plenty of AI algorithms such as Deep Learning involved in log anomaly detection, all methods without key words performs not good enough.

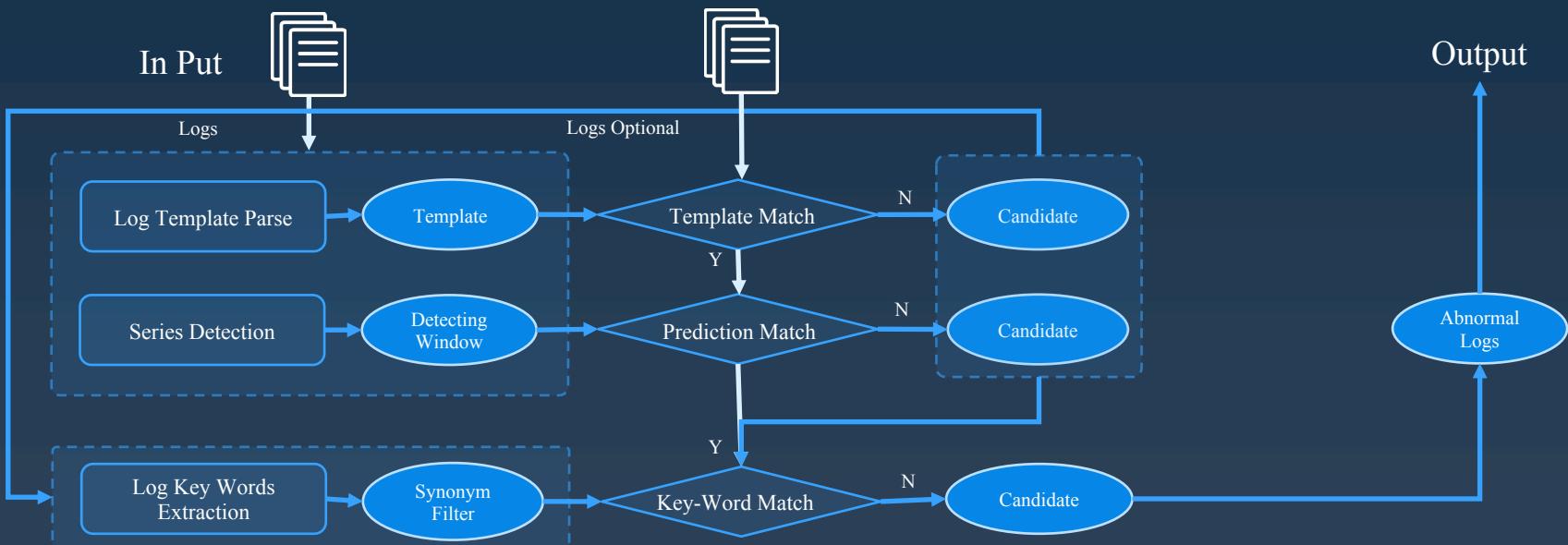
1. RISTO V, PIHELGAS M. LogCluster — a data clustering and pattern mining algorithm for event logs[C]// Conference on Net-work and Service Management (CNSM). 2015: 1-7.
2. HEPJ, ZHUJM, HESL, et al. Towards automated log parsing for large-scale log data analysis[J]. IEEE Transactions on Dependable and Secure Computing, 2018, 15(6): 931-944.
3. STUDIAWAN H, SOHEL F, PAYNE C. Automatic event log abstraction to support forensic investigation[C]// ACSW 2020. 2020: 1-9.
4. STUDIAWAN H, PAYNE C, SOHEL F. Automatic graph-based clustering for security logs[C]// Advanced Information Networking and Applications(AINA). 2019: 914-926.
5. DAI H, LI H, CHEN C S, et al. Logram: efficient log parsing using n-gram dictionaries[R]. 2020.
6. DU M, LI F F. Spell: online streaming parsing of large unstructured system logs[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(11): 2213-2227.



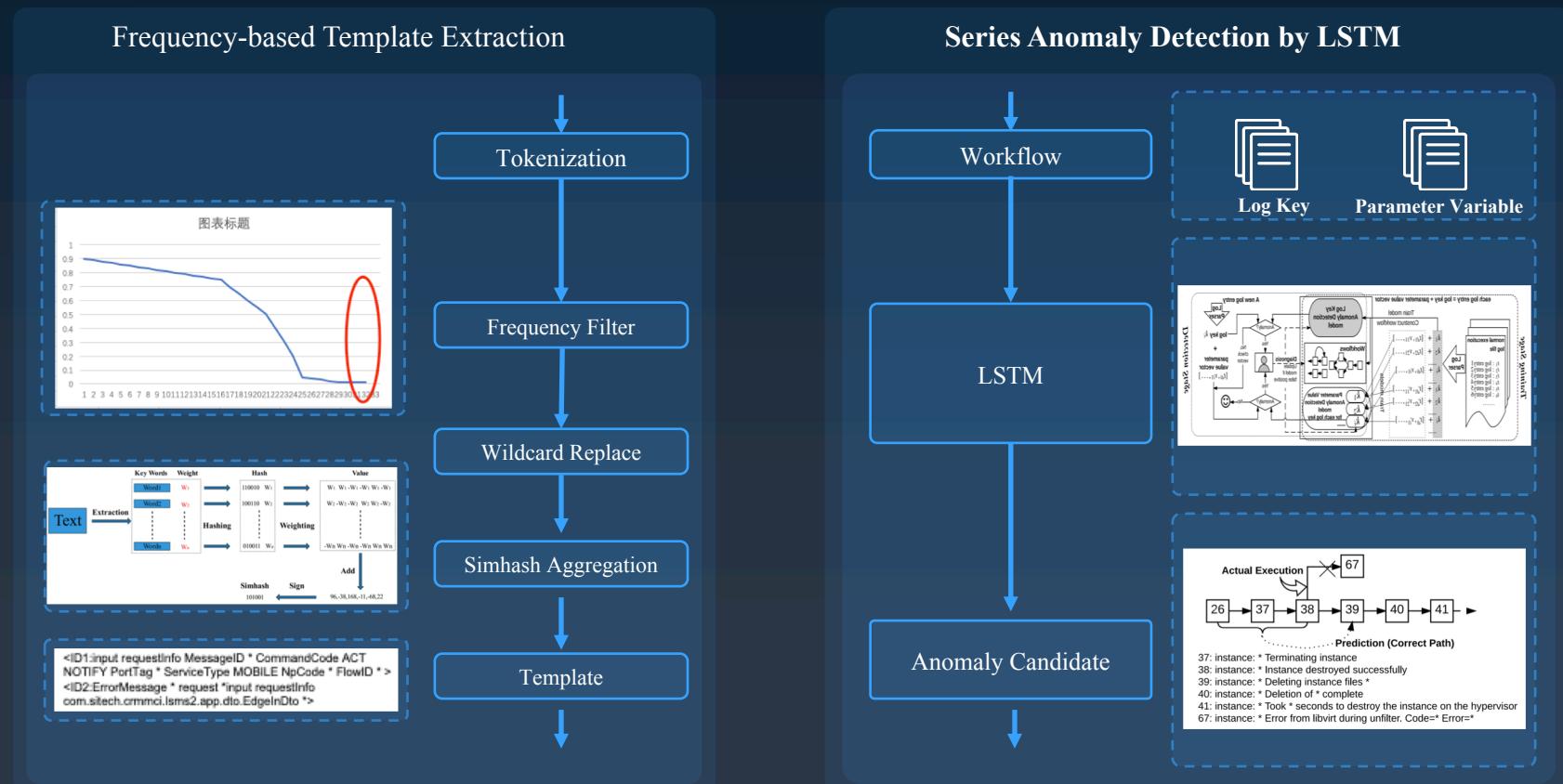
# Industry-level solution is empowered by evolution



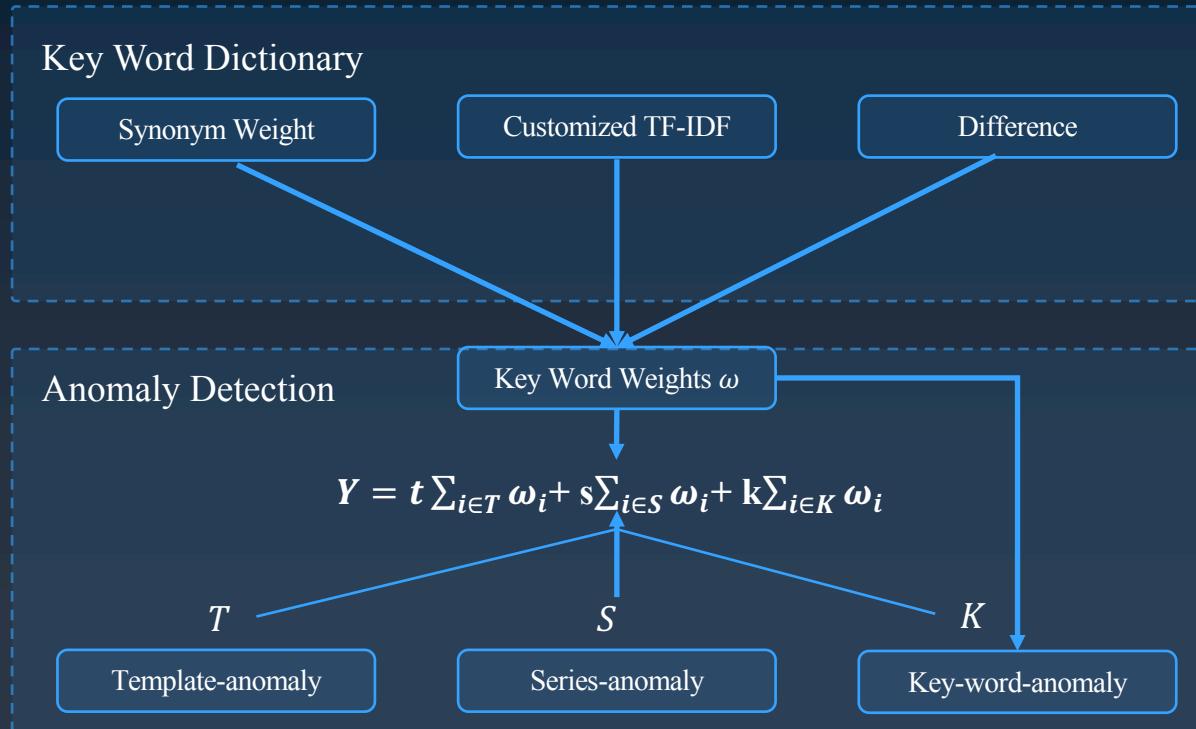
# Pre-detection is used to update key word dictionary for detection fixing



**Fast template extraction is utilized to make pre-detection based on LSTM much more efficient**



# Weighted detection with key-words fix generates the final anomaly



# The only 100 out of 100 score in competition dataset

Precision=100% ; Recall=100% ; **F1=1.00**

**Top 1/379** at China Unicom 5G+AI Network Application Special Competition .

- Advantages
  - ① More efficient template extraction ;
  - ② Increasingly evolved knowledge dictionary and potential knowledge graph construction ;
  - ③ Perfect tradeoff between time consumption and performance ;
  - ④ Optional fast detector with dictionary

## • Result



# Industrial Application is perfect

- Productization in AIOps
- Deployed for more than 20 systems involved in both BSS/OSS
- Daily processed 0.92TB data with 92% accuracy

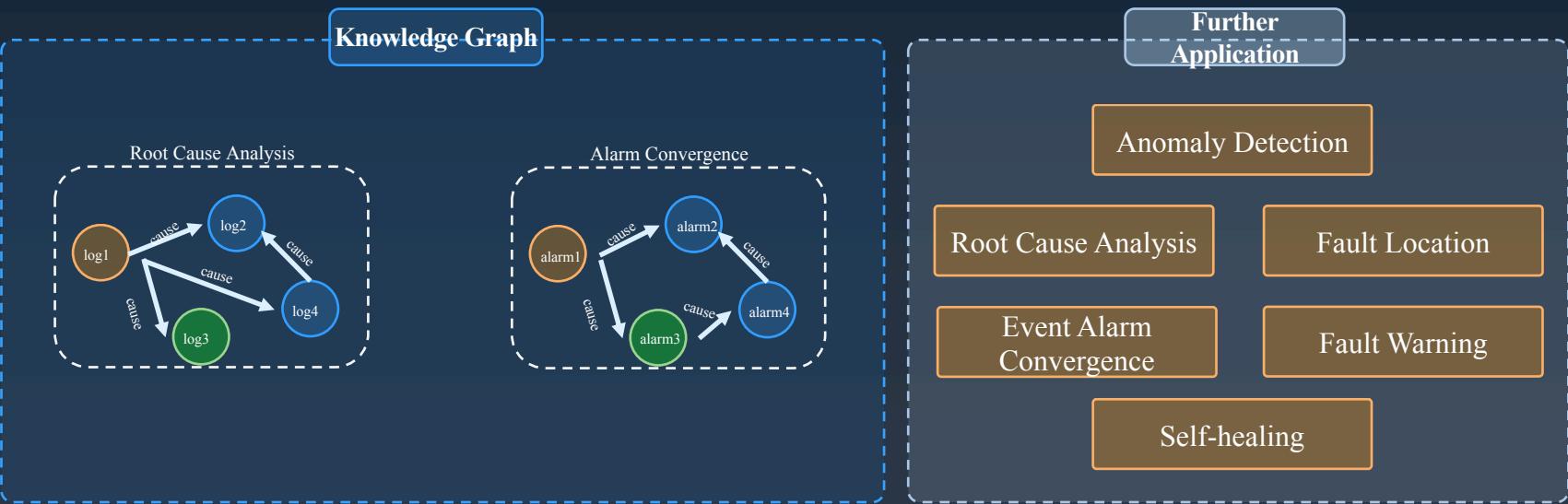
The image displays two side-by-side screenshots of a monitoring application interface, likely for an industrial system. Both screens have a header with tabs: '日志异常检测' (Log Anomaly Detection), '日志异常告警' (Log Anomaly Alert), and '检测策略配置' (Detection Strategy Configuration). The left screen shows a histogram of log anomalies over time (2021-11-20 00:00:00 to 2021-11-20 23:00:00) with a peak around 10 anomalies. Below the histogram is a table of detected anomalies:

异常描述	异常级别	文件名称	日志类型	日志对象	异常类型	模块内容	异常描述	异常时间	操作
1	严重	日志异常检测...	LomsSso	o.s.d.w.s...	新微式	--	【超上限异常 (15.00)】：检测模块日志量19...	2021-11-20 17:10:00	明细
2	严重	日志异常检测...	LomsSso	o.s.d.w.s...	新微式	--	【超上限异常 (15.00)】：检测模块日志量20...	2021-11-20 17:40:00	明细
3	严重	日志异常检测...	LomsSso	o.s.d.w.s...	新微式	--	【超上限异常 (15.00)】：检测模块日志量23...	2021-11-20 17:50:00	明细
4	严重	日志异常检测...	LomsSso	c.s.c.l.a.c...	新微式	--	【超上限异常 (15.00)】：检测模块日志量18...	2021-11-20 19:40:00	明细
5	严重	日志异常检测...	LomsSso	c.s.c.l.a.c...	模块日志量	反馈集群的应答 input soap...	【超上限异常 (46.434)】：检测模块日志量1...	2021-11-20 11:10:00	明细
6	严重	日志异常检测...	LomsSso	c.s.c.l.a.c...	模块日志量	接收CRM的请求 input rest...	【超上限异常 (22.357)】：检测模块日志量2...	2021-11-20 17:50:00	明细

The right screen shows a similar histogram and table for the 'c.s.c.l.a.c.CallService' module, with a single anomaly detected at 2021-11-20 18:50:00.

# Network will be more intelligent thanks to log anomaly detection

## Knowledge Graph Root Cause Analysis



# Thank you

## Q & A