

Assignment Report: literature review Deep Learning for Web Application Log Anomaly and Detection

Reviewer : Hesham Bashatah

Table of Contents

- Summary of Research Papers
- Critical Analysis
- Connections
- Conflicts Between the Findings
- Synthesis Matrix for Literature Review
- References

Summary of Research Papers:

Paper 1: DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning (Du et al., 2017)

The authors of DeepLog propose a model using the LSTM neural network approach that views the log messages almost as a sequence of events, similar to the words of a sentence. The model detects the next log message and highlight the ones that are unexpected as anomalies. DeepLog performs much better than traditional rule-based and statistical approaches to anomaly detection. The issue is that DeepLog focuses only on the anomaly detection but doesn't challenge any prediction and doesn't consider any performance metrics such as response time and system availability.

Paper 2: LogAnomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs (Meng et al., 2019)

This paper extends DeepLog by introducing Template2Vec embeddings -representation learning approach for system log templates- and an LSTM framework that captures both sequential and quantitative anomalies. The unsupervised method performs well without requiring labeled data and it handles unseen log patterns effectively. Evaluations on public datasets show higher precision and lower false-positive rates. However, it mainly targets system infrastructure logs and not focus on application-level factors like user experience or response time.

Paper 3: LogBERT: Log Anomaly Detection via BERT (Guo et al., 2021)

The authors propose a model called LogBERT (Bidirectional Encoder Representations from Transformers) that uses the idea of (masked log prediction) and teaches the model what is normal for log sequences. It far surpasses conventional RNN and LSTM models when dealing with long-range dependencies. It is good with any dataset but it is only using detection and does not look at what's going to come next but it can be used as a first phase of prediction.

Paper 4: Deep Learning for Anomaly Detection in Log Data: A Survey (Landauer et al., 2022)

This paper is a literature review that includes all the ways of doing log anomaly detection using deep learning including CNNs, RNNs, autoencoders, and even transformers. The authors describe the biggest challenges in web logs models such as messy data, imbalanced datasets, and models that are not very explainable and the difficulties of using such systems in the real world. They also recognize the challenges of doing log data and performance data combination. The paper doesn't involve any experiments.

Paper 5: BERT-Log: Anomaly Detection for System Logs Based on Pre-Trained Language Models (Chen et al., 2022)

BERT-Log utilizes the transfer learning capabilities of pre-trained NLP models to capture the semantics of log statements automatically so that no need for manual feature engineering. Experiments demonstrate superior performance compared to traditional methods and DeepLog but the model remains computationally intensive and demands substantial training resources.

Paper 6: Practitioner Survey on Log Anomaly Detection (2024)

This paper presents a comprehensive survey of over 300 industry practitioners, aiming to investigate the primary challenges associated with the implementation of log anomaly detection in operational settings. The findings indicate that the most significant obstacles relate to model explainability, data privacy concerns, and the instability of models when log structures are modified. The study underscores the discrepancy between academic research and industrial practice noting that while research models often demonstrate strong performance in controlled environments but they frequently lack the robustness, interpretability, and adaptability required for production deployment. The paper provides a valuable empirical perspective on the practical limitations and unmet needs within the field.

Critical Analysis

Recent studies show that deep learning techniques have significantly improved how anomalies are detected in log data. Models based on LSTMs and Transformers have demonstrated strong capabilities in identifying abnormal behaviors and patterns. However, these models are primarily effective at detection not at predicting future issues and they tend to perform best with structured data. This is a concern because log data in production systems particularly in web applications is often noisy, inconsistent, and unstructured.

Although deep learning models deliver noticeable gains in accuracy but they still face challenges related to interpretability, scalability, and their ability to adapt to different log formats. Modern architectures such as LogBERT and BERT-Log outperform traditional approaches but are resource intensive and require substantial computational power. The findings from the practitioner survey further highlight a clear gap between research advancements and industry requirements emphasizing the ongoing need for solutions that are more explainable, scalable, and easier to deploy in production environments.

Connections

1. **Focus on deep models:** Each of the papers focuses on employing deep models such as LSTM, BERT, and Autoencoders for anomaly detection within the log data.
2. **Focus on accuracy of detection:** They are all attempting to improve the accuracy of detection and reduce false positives.
3. **Easy comparison:** Most of these use HDFS and BGL datasets, so comparison of results is easy.
4. Understanding when and where log events occur is key to making sense of system behavior.
5. **They all understand the tough challenges:** messy log parsing, feature engineering, and changing logs.

Conflicts Between Findings

1. **Supervised vs. unsupervised approaches:** Some rely on labeled data i.e. DeepLog while others i.e. LogAnomaly learn patterns without supervision.
2. **Model complexity:** Simpler RNNs are faster whereas transformers achieve higher accuracy but require more resources.
3. **Detection vs. prediction:** Existing works mainly detect anomalies but few aim to forecast them proactively.
4. **Dataset dependency:** Results vary depending on whether structured HDFS or raw logs are used.

Synthesis Matrix for Literature Review

Study	Authors	Year	Research Design	Key Findings / Results	Methodology	Limitations
DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning	Min Du, Feifei Li, Guineng Zheng, Vivek Srikumar	2017	Experimental	Achieves high anomaly detection accuracy Learns log sequence patterns Reduces false alarms	LSTM-based sequence prediction model treating logs as event sequences	Focuses on detection only Cannot predict future incidents Relies on structured log templates
LogAnomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs	Weibin Meng ,Ying Liu ,Yichen Zhu ,Shenglin Zhang ,Dan Pei ,Yuqing Liu ,Yihao Chen ,Ruizhi Zhang ,Shimin Tao ,Pei Sun ,Rong Zhou	2019	Experimental	Improves detection of unseen/novel log patterns Effective without labels	Template2Vec embeddings + LSTM anomaly detection	Limited to system logs Does not handle performance metrics (latency, availability) Not predictive
LogBERT: Log Anomaly Detection via BERT	Haixuan Guo, Shuhan Yuan, Xintao Wu	2021	Self-supervised learning	Outperforms LSTM-based methods Strong contextual understanding Robust to template changes	Transformer-based BERT model trained using self-supervised masked token prediction	Very computationally expensive Focuses on detection not forecasting Heavy preprocessing required

Deep Learning for Anomaly Detection in Log Data: A Survey	M. Landauer et al., Max Landauer, Sebastian Onder, Florian Skopik, Markus Wurzenberger	2022	Systematic literature review	Maps deep learning methods for log anomaly detection Identifies gaps in interpretability & deployment	Qualitative review of DL approaches (CNN, RNN, Autoencoders, Transformers)	No experimental validation Cannot compare models under identical conditions
BERT-Log: Anomaly Detection for System Logs Based on Pre-trained Language Models	Song Chen, Hai Liao	2022	Experimental	Achieves higher F1-scores than traditional models Improves semantic capture of log events	BERT fine-tuning on log templates Transfer learning from NLP	High compute cost Domain specific adaptation required Detection focused only
Practitioner Survey: Expectations on Log Anomaly Detection	Xiaoxue Ma, Yishu Li, Jacky Keung, Xiao Yu, Huiqi Zou, Zhen Yang, Federica Sarro and Earl T. Barr	2024	Mixed method survey (questionnaires & interviews)	Shows disconnect between academic methods and industry needs (explainability, robustness, real time performance)	Survey based research on practitioner perspectives	No model proposed No quantitative experiments Insights depend heavily on participant experience

References

1. Min Du, Feifei Li, Guineng Zheng, Vivek Srikumar, 2017. DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning. <https://users.cs.utah.edu/~lifeifei/papers/deeplog.pdf>
2. Weibin Meng ,Ying Liu ,Yichen Zhu ,Shenglin Zhang ,Dan Pei ,Yuqing Liu ,Yihao Chen ,Ruizhi Zhang ,Shimin Tao ,Pei Sun ,Rong Zhou , 2019. LogAnomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs. <https://www.ijcai.org/proceedings/2019/0658.pdf>
3. Haixuan Guo, Shuhan Yuan, Xintao Wu, 2021. LogBERT: Log Anomaly Detection via BERT. <https://arxiv.org/abs/2103.04475>
4. M. Landauer et al., Max Landauer, Sebastian Onder, Florian Skopik, Markus Wurzenberger, 2022. Deep Learning for Anomaly Detection in Log Data: A Survey. <https://arxiv.org/abs/2207.03820>
5. Song Chen, Hai Liao, 2022. BERT-Log: Anomaly Detection for System Logs Based on Pre-Trained Language Models. <https://www.tandfonline.com/doi/full/10.1080/08839514.2022.2145642>
6. Xiaoxue Ma, Yishu Li, Jacky Keung, Xiao Yu, Huiqi Zou, Zhen Yang, Federica Sarro and Earl T. Barr, 2024. Practitioner Survey on Log Anomaly Detection. <https://arxiv.org/html/2412.01066v1>