

Retrieval-Augmented Generation and Chain-of-Thought Prompting Robustness to Distribution Shifts in Zero-Shot Fault Detection

Assignee Research

June 4, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How do retrieval-augmented generation and chain-of-thought prompting differ in their robustness to distribution shifts when applied to zero-shot fault detection in cyber-physical systems using. 16 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: A Big Data Architecture Design for Smart Grids Based on Random Matrix Theory. Research question: How do retrieval-augmented generation and chain-of-thought prompting differ in their robustness to distribution shifts when applied to zero-shot fault detection in cyber-physical systems using open-weight LLMs?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.5/10.

3 Results

16 papers retrieved. 16 claims extracted; 2 independently verified. Quality review score: 4.5/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
Imprecise formulas, biased sample selections, and improper training processes lead to unsatisfied results as system size	×	0.02
Data-driven analysis tools are more suitable than model-based tools for complex large-scale interconnected systems with	✓	0.16
Traditional model-based tools cannot handle 4Vs data (volume, variety, velocity, veracity) in smart grids within tolerab	×	0.11
Big data technology has been successfully applied to quantum systems, financial systems, biological systems, and wireles	×	0.04
In the provided notation, 'hat' denotes raw data and 'tilde' denotes transformation data formed by normalization.	×	0.04
The covariance matrix S is defined as $S = (1/N) * X * X^H$ in $C N\$ \times \N space.	×	0.08
The parameter c is defined as the ratio of rows to columns ($c = N/T$).	×	0.03
In the experimental setup shown in Table (p3), the sampling time (T_{samp}) is 1000, N is 400, and n is 1000.	×	0.04
The parameters for the Marcenko-Pastur Law distribution in the experiment are $a=0.13509$, $b=2.6649$, and $h=0.1$.	×	0.02
PBus-n represents the power demand of load at bus-n.	×	0.02
γ_{acc} and γ represent the addition and multiplication factors for load fluctuations of the grid, respectively.	×	0.02
A PQ Node is characterized by 2n given numbers (P, Q) and 2n unknown numbers (V, θ).	×	0.03
The proposed architecture allows regional Mean Spectral Radius (MSR) calculation using only local measured or simulated	✓	0.27
Comparative analysis of distributed MSRs produces a contour line capable of detecting events and locating their source.	×	0.12
Five case studies and their visualizations were conducted to validate the designed architecture.	×	0.03
This study claims to be the first attempt to apply big data technology into smart grids.	×	0.12

References

- <http://arxiv.org/abs/1501.07329v4>
- <http://arxiv.org/abs/2309.02144v1>
- <http://arxiv.org/abs/1910.14127v2>