Understanding Primary Fault Types of IT Failures

An Analysis of Public Incident Reports Using Large Language Models

1 Introduction

- Modern organisations increasingly operate as software-defined businesses, relying on fast-paced agile methodologies [1, 2].
- This transformation, while beneficial, raises the risk of operational incidents, e.g. unplanned service interruptions or degradations.
- Research shows 70% of outages stem from changes to live systems, highlighting the need for effective incident analysis [3].
- AIOps promises improvement, but is limited by inconsistent, manual reporting [4, 5].
- Rise of cloud-native architectures introduces new fault patterns.
- Lack of cross-organizational analysis on fault types, severity, and evolution.

Research Questions:

1. What taxonomy of primary fault types can be established for modern IT incidents, and how reliable is automated classification of these categories?
2. What is the relative frequency of different primary fault types across incident reports?
3. Are there correlations between specific primary fault types and the duration of incidents?
4. Has the frequency of specific fault types changed over time, particularly with the adoption of cloud-native architectures?

2 Methodology

- Scrape incident reports from the Verica Open Incident Database API [16]
- Filter out incomplete reports with missing descriptions or duration data, exclude years with not enough reports, resulting in 7,804 incident reports from years 2014 to 2022
- Classify incident reports using the Athene V2 72B large language model (LLM) [18] on the Delft Blue supercomputing infrastructure [19]
- Evaluate model's performance using a manually labelled ground truth, sampled via a population-proportional stratified design to ensure coverage across fault types.
- Use Scikit-learn, Pandas, and NumPy to measure fault category frequencies, analyze correlations with incident duration, and evaluate time trends.
- Develop clear, insightful visualizations using Matplotlib and Seaborn, e.g. bar charts of fault type distributions, line plots of temporal changes

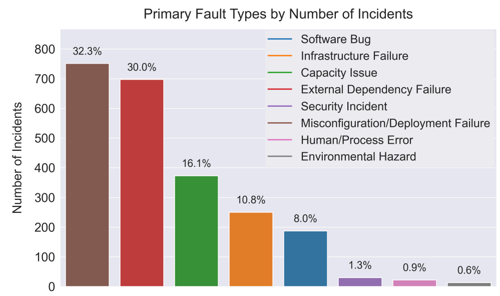
3 Results

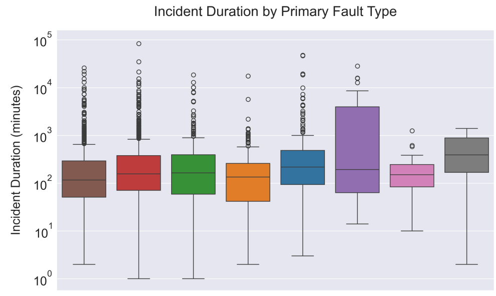
Taxonomy of Primary Faults and Classification Performance

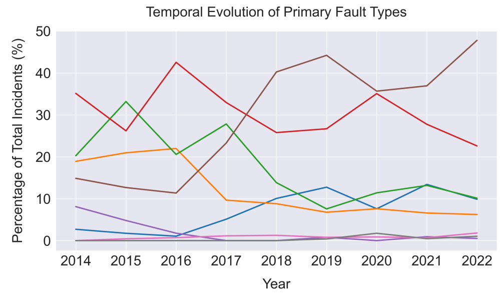
Technical	Human/Organizational	External/Environmental	Other
Infrastructure Failure Software Bug Misconfiguration/Deployment Failure Capacity Issue	Human/Process Error *(operator mistakes, inadequate procedures, monitoring/alerting failures, etc.)*	External Dependency Failure Security Incident Environmental Hazard *(power, fire, fiber cut, etc.)*	Scheduled Maintenance Unknown

Overall metrics: Accuracy: 92%, Macro F1-score: 0.89

Analysis of Primary Fault Types Frequency, Severity, and Evolution







Upward trend ↑
Software Bug, Misconfiguration/Deployment Failure, Environmental Hazard

Downward trend ↓
Infrastructure Failure, Capacity Issue

No trend —
External Dependency Failure, Security Incident, Human/Process Error

4 Conclusions

- Automation improves speed but introduces new risks, configuration management is critical.
- External dependencies need stronger resilience strategies, shifting from traditional infrastructure concerns, as they extend beyond organizational boundaries.
- Rise of software bugs calls for more testing, security incidents are rare but highly disruptive.
- Cloud technologies successfully reduce infrastructure failures and capacity issues.
- LLM-based classification is viable for large-scale analysis, but can be improved via fine-tuning.

TU Delft Delft University of Technology

Author: Jakub Rutkowski (j.a.rutkowski@student.tudelft.nl)
Supervisor: Eileen Kapel
Responsible Professor: Diomidis Spinellis

Scan for references, paper, and replication package