REAL TIME ANOMALY DETECTION FOR CONTAINERIZED ENVIRONMENTS

Nirothipan Megalingham
219375K

Degree of MSc in Computer Science

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2024

REAL TIME ANOMALY DETECTION FOR CONTAINERIZED ENVIRONMENTS

Nirothipan Megalingham
219375K

Thesis submitted in partial fulfillment of the requirements for the degree
Degree of MSc in Computer Science

Department of Computer Science and Engineering
Faculty of Engineering

University of Moratuwa
Sri Lanka

June 2024

DECLARATION

I declare that this is my own work and this Thesis does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date: 01/09/2024

The supervisor should certify the Thesis with the following declaration.

The above candidate has carried out research for the Degree of MSc in Computer Science Thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Dr. Sunimal Rathnayake

Signature of the Supervisor: Date: 23/09/2024

ACKNOWLEDGEMENT

I extend my deepest and most sincere gratitude to Dr. Sunimal Rathnayake of the University of Moratuwa, Faculty of Computer Science and Engineering, for his invaluable mentorship during this research endeavor. Without his unwavering support, navigating through this journey would have been nearly insurmountable. Dr. Rathnayake not only provided continuous guidance but also mentored me during the thesis passing the time zone barriers, offering assistance remotely and helped me maintain the momentum during the challenging phases.

A heartfelt thank you is also due to my family for their enduring patience and encouragement throughout my research program. Their unwavering support has been a constant source of strength, from the inception of my studies to the culmination of this endeavor. Furthermore, I express my deepest appreciation to my parents for their steadfast support throughout my academic journey.

I am also indebted to the University of Moratuwa for granting me the opportunity to participate in the MSc program and for furnishing the requisite resources to see this research to fruition. Additionally, I am grateful to my former workplace, WSO2, for accommodating my pursuit of this part-time MSc and research alongside my professional responsibilities. Their support has been indispensable in achieving this milestone.

ABSTRACT

Anomalies in containerized environments pose a significant threat, given their potential to escalate small failures into catastrophic outcomes. These anomalies, which can manifest in various forms, possess the capability to disrupt service level agreements and tarnish an organization's reputation irreversibly. Thus, it becomes imperative to detect and address these anomalies promptly to minimize their adverse effects on business operations. In this study, we delve into the landscape of anomalies prevalent in containerized environments, focusing on understanding their diverse nature and the substantial impact they can have.

In this comprehensive survey, our focus lies in the real-time detection of anomalies within Kubernetes environments, a critical aspect in ensuring the robustness and stability of modern containerized systems. To achieve this, we conducted an extensive literature review, delving into existing research and methodologies pertinent to anomaly detection within Kubernetes ecosystems.

Furthermore, we conducted practical experiments by deploying applications on Azure Kubernetes Services, leveraging the inherent Kubernetes metrics API and Prometheus API to gather pertinent data. Employing sophisticated feature selection techniques, we curated datasets and trained a decision tree model capable of discerning anomalous patterns in real-time metrics streams. Through rigorous experimentation, we validated the efficacy of our approach, achieving a remarkable accuracy rate of 95% in anomaly prediction. This research underscores the significance of real-time anomaly detection in Kubernetes environments and offers tangible insights for enhancing the resilience of containerized infrastructures.

Keywords: Real Time Anomaly Detection, Machine Learning, Containerization, Kubernetes

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
AKS	Azure Kubernetes Service
CICD	Continuous Integration and Continuous Deployment
HHMM	Hierarchical Hidden Markov Models
K8s	Kubernetes
LSTM	Long short-term memory
MSA	Microservice Architecture
NFV	Network Function Virtualization
S3	Amazon Simple Storage Service
VM	Virtual Machines

LIST OF APPENDICES