**CONGESTION CONTROL PREDICTION MODEL**

**FOR 5G ENVIRONMENT**

**BASED ON**

**SUPERVISED AND UNSUPERVISED MACHINE LEARNING APPROACH**

*A Seminar Report*

*Submitted to the APJ Abdul Kalam Technological University in*

*partial fulfillment of the requirements for the award of degree*

***Bachelor of Technology***

*in*

***Computer Science and Engineering***

By

**JISSIN SAM MATHEW**

**PTA21CS035**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**COLLEGE OF ENGINEERING KALLOOPPARA**

**KERALA**

**October 2024**

**DEPARTMENT OF COMPUTER SCIENCE AND**

**ENGINEERING**

**COLLEGE OF ENGINEERING KALLOOPPARA**

**2024-25**



**CERTIFICATE**

This is to certify that the report entitled **Congestion Control Prediction Model for 5G Environment Based on Supervised and Unsupervised Machine Learning Approach** submitted by **JISSIN SAM MATHEW (Reg. no.PTA21CS035)**,to the APJ Abdul Kalam Technological University in partial fulfilment of the B.Tech degree in Computer Science & Engineering is a bonafide record of the seminar work carried out by him under our supervision.This report in any form has not been submitted to any other University or Institution for any purpose.

|  |  |  |
| --- | --- | --- |
| **Mrs. Shilpa Sajeev**  **Internal Guide** | **Mrs. Anitha Jose**  **Coordinator** | **Dr. Renu George**  **Head of the Department** |
| Department Of Computer  Science & Engineering | Department Of Computer  Science & Engineering | Department Of Computer  Science & Engineering |

**ABSTRACT**

With the emergence of 5G technology, congestion control has become a vital challenge to be addressed in order to have efficient communication. There are several congestion control models that have been proposed to control and predict the possible congestion in 5G technology. However, finding the optimal congestion control model is an important yet challenging task. In this paper, we examine the supervised and unsupervised machine learning approaches to the task of predicting the possible node that causes congestion in the 5G environment. Due to the huge variance in the domains of the data set columns, measuring the prediction’s consistency was not an easy task. During our study, we tested twenty-six supervised and seven clustering algorithms. Finally, and based on the performance criteria, we have identified the best five algorithms out of the studied algorithms.

**ACKNOWLEDGEMENT**

I take this opportunity to express my deepest sense of gratitude and sincere thanks to everyone who helped me to complete this work successfully. I express my sincere thanks to **Dr. Nisha Kuruvilla** (Principal) for the constant support and help. I also thank **Dr. Renu George**(Head of Department) Computer Science and Engineering, College of Engineering Kallooppara for providing me with all the necessary facilities and support. I would like to express my sincere gratitude to **Mrs. Anitha Jose, Mrs. Maniveena, and Mrs. Shilpa Sajeev**, department of Computer Science and Engineering,College of Engineering Kallooppara for their support and cooperation. I would like to place on record my sincere gratitude to my seminar guide **Mrs.Shilpa Sajeev**, Assistant Professor, Computer Science and Engineering, College of Engineering Kallooppara for the guidance and mentorship throughout the course. Finally, I thank my family, and friends who contributed to the successful fulfilment of this seminar work.

**JISSIN SAM MATHEW**

**CONTENTS**

**LIST OF FIGURES**

**LIST OF TABLES**

**LIST OF ABBREVIATIONS**

**CHAPTER 1**

**INTRODUCTION**

* 1. **BACKGROUND**

5G networks, compared to previous generations, offer higher speeds, lower latency, and improved coverage. These advantages have led to their widespread adoption [1]. However, with the growing number of nodes connected to the network, new challenges have emerged, particularly in the area of congestion control [2]. The objective of routing algorithms is to select the best path while avoiding congestion, but this can lead to increased routing costs [3]. Congestion can cause significant delays and reduced throughput, making it a critical issue during 5G routing decisions.

Several approaches have been explored to manage congestion in the 5G environment [4], [5]. One feature that helps reduce congestion in 5G networks is the ability to dynamically distribute resources such as bandwidth and frequencies. Techniques like network slicing [6] and edge computing [7] enable traffic-based optimization in 5G networks, reducing congestion. Moreover, implementing Quality of Service (QoS) ensures that vital services, like emergency communications, remain unaffected by current traffic, as such traffic is given higher priority, and resources are allocated accordingly [8]. Another method of managing congestion is traffic offloading, where data is transferred to alternative networks, such as Wi-Fi [9], thereby reducing the load on the 5G infrastructure.

Machine learning (ML) algorithms have also been applied to manage network congestion, with promising results [10]. Supervised ML algorithms are trained using labeled data, while unsupervised algorithms utilize unlabeled data [11]. Both types of algorithms are used to predict potential congestion points and to determine the optimal congestion control window. In supervised ML, classification plays a critical role by grouping data into predefined classes. In contrast, unsupervised ML uses clustering to categorize data into groups without prior class labels. The choice between supervised

and unsupervised algorithms depends on factors like the size and complexity of the data, as well as the required accuracy.

Machine learning plays a crucial role in addressing congestion control challenges within 5G networks. Selecting the most suitable algorithm for predicting congestion is essential to efficiently identify congestion nodes and optimize network performance. Various machine learning approaches have been explored in this area, including the evaluation of 26 supervised and 7 unsupervised algorithms. Unsupervised learning methods, such as clustering, have demonstrated effectiveness in improving congestion prediction accuracy by grouping similar data into clusters. To ensure the reliability of the models, consistency was assessed using Cronbach’s alpha, a metric for evaluating the internal consistency of the datasets. Additionally, the performance of the algorithms was measured using key metrics such as True Positive (TP) and False Positive (FP) rates, precision, recall, the Receiver Operating Characteristic (ROC) curve, and the Area Under the Curve (AUC).

* 1. **PURPOSE AND SCOPE**

The Congestion Control Prediction Model for 5G networks, using both supervised and unsupervised machine learning (ML) approaches, aims to enhance the efficiency and reliability of 5G networks by accurately predicting congestion nodes and optimizing network routing. This model seeks to address the growing complexity in managing network traffic in 5G environments, which are characterized by high data volumes, increased node density, and latency-sensitive applications.

Key aspects of the purpose and scope include:

**Congestion Prediction Accuracy:** By leveraging supervised ML algorithms for classification and unsupervised techniques such as clustering, the model enhances the ability to predict network congestion more accurately, minimizing delays and improving throughput.

**Optimized Resource Allocation:** The model enables 5G networks to dynamically allocate resources like bandwidth and frequencies, ensuring efficient use of available network resources and preventing bottlenecks that could degrade performance.

**Real-time Decision Making:** The integration of ML allows for real-time analysis of network conditions, ensuring that congestion control measures are applied proactively, reducing the chances of critical delays, especially in high-priority applications like emergency services or autonomous vehicles.

**Scalability and Adaptability:** The model is designed to scale with the growing size and complexity of 5G networks. Its ability to handle large data sets and diverse node architectures ensures that it can adapt to changing network conditions and maintain optimal performance.

**Performance Metrics:** The model evaluates algorithms based on performance metrics such as True Positive (TP) and False Positive (FP) rates, precision, recall, and Area Under the Curve (AUC). These metrics ensure that the chosen ML algorithms provide reliable congestion control predictions and contribute to better overall network health.

**Machine Learning-Driven Optimization:** Both supervised and unsupervised ML techniques are employed to continuously improve the accuracy of congestion predictions. This data-driven approach ensures that the model adapts to the dynamic nature of 5G traffic, providing better congestion mitigation over time.

In summary, the purpose of this congestion control prediction model is to provide a more intelligent and scalable approach to managing network congestion in 5G environments. By leveraging supervised and unsupervised machine learning techniques, the model predicts congestion points in real-time, enabling proactive measures to optimize traffic flow. This reduces delays, packet loss, and ensures priority for critical services like emergency communications and low-latency applications.

The scope covers a wide range of 5G applications, including bandwidth-heavy services and latency-sensitive use cases. The model is adaptable to different 5G deployments, ensuring optimal network performance and maintaining high quality of service in dynamic environments.

**CHAPTER 2**

**LITERATURE SURVEY**

**2.1 A Heuristic Path Search for Congestion Control in Wireless Sensor Networks [13]:**By G. Sangeetha, M. Vijayalakshmi, S. Ganapathy, and A. Kannan proposed a heuristic model aimed at addressing congestion in Wireless Sensor Networks (WSNs) by focusing on data loss and energy reduction. The model involves regular adjustments to the sensor nodes' topology based on node degree and time intervals. This approach enhances power consumption management and reduces interference, thereby providing a more effective energy-congestion-aware technique for routing, termed Survival Path Routing (SPR). SPR is particularly beneficial for IoT applications operating in high-traffic environments, where multiple nodes attempt to send packets to destination nodes simultaneously.

**2.2 Congestion-Aware Opportunistic Routing Protocol in Wireless Sensor Networks by M. Shelke, A. Malhotra, and P. N. Mahalle[19]:** Shelke et al. proposed a novel routing algorithm aimed at selecting the most efficient routes in wireless sensor networks (WSNs) by integrating adaptive sleep scheduling mechanisms grounded in opportunistic theory. This innovative approach focuses on minimizing energy consumption and network congestion by intelligently managing the active and inactive states of sensor nodes. By evaluating multiple potential routes and dynamically adjusting to varying network conditions, the protocol enhances data transmission efficiency while reducing packet loss and energy depletion. The implementation of this congestion-aware routing strategy not only ensures better resource utilization but also prolongs the lifespan of the sensor network, making it particularly effective in high-traffic scenarios where numerous nodes attempt simultaneous data transmission. The findings underscore the significance of adaptive routing mechanisms in optimizing network performance and highlight the potential for improved management in congestion-prone environments.

**2.3 Improving TCP congestion control with machine intelligence by Kong, Zang, and Ma** [58]:Kong, Zang, and Ma developed dual machine-learning approaches to address TCP congestion control issues in under-buffered connections over wired environments. Their model incorporated supportive and adaptive loss prediction mechanisms to optimize the trade-off between delay and throughput. By effectively predicting packet loss, the approach enhanced the responsiveness of TCP to network conditions, leading to improved data transmission rates and overall network performance. The integration of machine learning allowed for dynamic adjustments to congestion control parameters, demonstrating significant advancements in managing network congestion and ensuring reliable communication in challenging environments.

**2.4 Centralized and localized data congestion control strategy for vehicular ad hoc networks using a machine learning clustering algorithm** [59]:In their research, Taherkhani and Pierre employed the K-means algorithm to effectively manage congestion within Vehicle Ad-hoc Networks (VANETs). The approach is structured into three main components: directing traffic flows, detecting instances of data congestion, and clustering communications for efficient resource allocation. This multi-faceted strategy not only enhances communication efficiency but also aids in reducing overall network congestion by optimizing data transmission paths, ensuring smoother vehicular communication and improved network performance.

**2.5 A Machine Learning Approach to Improve Congestion Control Over Wireless Computer Networks by Geurts, El Khayat, and Leduc [34]:** Geurts et al. proposed a machine learning model that utilizes an automatic loss classifier, leveraging a simulated database of random network topologies to enhance congestion control in wireless computer. By analyzing diverse network configurations and traffic patterns, the model effectively learns to classify conditions that lead to congestion, enabling proactive management to mitigate potential issues. The research highlights the significant role of machine learning in creating adaptive, intelligent systems capable of optimizing network performance and ensuring robust communication in dynamic environments. This approach not only improves the overall efficiency of wireless networks but also contributes to the ongoing advancement of automated network management techniques.