Online at: https://researchtrendsjournal.com **ISSN No:** 2584-282X

INTERNATIONAL JOURNAL OF TRENDS IN EMERGING RESEARCH AND DEVELOPMENT

Volume 2; Issue 4; 2024; Page No. 170-175

Received: 03-05-2024
Accepted: 07-06-2024

5g network management aspects enhanced with machine learning

[1]Divyanshu Rao and [2]Dr. Shubham Dubey

[1]Research Scholar, Shri Krishna University Chhatarpur, Madhya Pradesh, India
[2]Assistant Professor, Shri Krishna University Chhatarpur, Madhya Pradesh, India

DOI: https://doi.org/10.5281/zenodo.14949069

Corresponding Author: Divyanshu Rao

Abstract

The aim was to analyze, propose, and improve the state-of-the-art techniques to bring 5G networks one step closer to what we envision it could be. Techniques such as D2D/IoT (Device-to-Device/Internet of Things), Mesh Networks, Ultra-densification, Dynamic channel selection, Network sharing, Network slicing, Multi-connectivity/Multi-link, Carrier Aggregation, and the utilization of new frequencies such as 3.5 GHz and narrow band were analyzed, improved, and presented, achieving good results in all tested cases. These proposals could be further tested in a real environment and formulated by the research community and standards to be integrated into the new versions of 3GPP for 5G networks and future generations.

Keywords: 5G, network, management, machine learning, number and development

Introduction

As the number of smartphones and demand for higher data rate connections continue to grow explosively, technology must evolve to keep pace and provide suitable communication schemes. It is estimated that the total global mobile wireless user devices will exceed 10 billion, with mobile data traffic growing more than 200 times compared to 2010 figures. The new 5G communication systems promise a complete network structure with unlimited access to information, providing service demands to users far beyond what 4G offers by supporting innovative new wireless technologies and network architecture to meet the extremely high-performance requirements. These features include support for new types and massive numbers of devices, very high mobile traffic volumes, universal access for users, very high frequency reuse and spectrum reuse in wireless technologies, automated provisioning, configuration, and management of a wide range of new network services, ultra-reliable, ultra-low latency, ultra-densification, and more.

5G networks would be heterogeneous networks (HetNets), meaning that different networks will be integrated together into a unified system, enabling aggregation of multiple existing radio access technologies (RATs) such as LTE-A, WiFi, D2D, and even lightly-licensed. Delivering all 5G requirements in order to support the new features and services, a substantial change in the network architecture is inevitable. Normally in the past, such a process would need deployment of specialized devices built for a specific application and with fixed functionalities. Thus, any development and transformation to follow the constantly increasing and heterogeneous market requirements demands a huge investment to change/deploy hardware. Nowadays, various technologies and architectures have been utilized in order to solve this problem and provide a faster introduction and adaptation of new technologies to the communication systems.

Literature Review

Singh, Shashank. (2023) [1]. The rapid evolution of 5G technology necessitates innovative approaches for network management to ensure reliability, security, and efficiency. This research paper investigates the application of Artificial Intelligence (AI) in 5G network management, offering a multi-dimensional evaluation of its impacts and potentials. Through an integrative methodology combining literature review, case studies, and simulations, the study identifies critical roles for AI in areas such as automated fault detection, resource allocation, security, and Quality of Service (QoS) management. Our findings indicate that AI algorithms significantly enhance network performance metrics, thus acting as a vital component for the optimization of 5G networks. This paper concludes by highlighting the broader operational and economic implications of integrating AI into 5G network management and suggests directions for future research.

Barakabitze, Alcardo *et al*. (2022) [2]. This chapter provides background information regarding the application of Machine Learning (ML) and artificial intelligence (AI) in managing big data streaming in future softwarized and virtualized 5G networks. It discusses the concepts for optimization of data management with ML in softwarized 5G networks. The chapter presents state-of-the-art research work on multimedia big data analytics by providing the current big data 5G frameworks and their applications in multimedia analyses. The learning paradigms in ML that influence the data collection, feature engineering, and the establishment of ground truth are divided into four categories, namely: supervised, unsupervised, semi supervised, and reinforcement learning. Multimedia big data is often multimodal, heterogeneous, and unstructured, attributes that make it difficult to represent and model. The deep learning models for Internet of Things (IoT) big data analytics can be performed at the IoT cloud layer on high-performance computing systems or cloud platforms.

Hernández-Chulde, Carlos *et al*. (2019) [3]. The adoption of Software Define Networking (SDN), Network Function Virtualization (NFV) and Machine Learning (ML) will play a key role in the control and management of 5G network slices to fulfill the specific requirements of application/services and the new requirements of fifth generation (5G) networks. In this research, we propose a distributed architecture to perform network analytics applying ML techniques in the context of network operation and control of 5G networks.

Kaur, Jasneet *et al*. (2021) [4]. Wireless communication systems play a very crucial role in modern society for entertainment, business, commercial, health and safety applications. These systems keep evolving from one generation to next generation and currently we are seeing deployment of fifth generation (5G) wireless systems around the world. Academics and industries are already discussing beyond 5G wireless systems which will be sixth generation (6G) of the evolution. One of the main and key components of 6G systems will be the use of Artificial Intelligence (AI) and Machine Learning (ML) for such wireless networks. Every component and building block of a wireless system that we currently are familiar with from our knowledge of wireless technologies up to 5G, such as physical, network and application layers, will involve one or another AI/ML techniques. This overview paper, presents an up-to-date review of future wireless system concepts such as 6G and role of ML techniques in these future wireless systems. In particular, we present a conceptual model for 6G and show the use and role of ML techniques in each layer of the model. We review some classical and contemporary ML techniques such as supervised and un-supervised learning, Reinforcement Learning (RL), Deep Learning (DL) and Federated Learning (FL) in the context of wireless communication systems. We conclude the paper with some future applications and research challenges in the area of ML and AI for 6G networks.

Khan, Imran *et al*. (2020) [5]. As the demand for faster and more reliable mobile networks intensifies, the deployment of 5G has emerged as a transformative solution to meet the growing needs of connectivity. However, to fully leverage the potential of 5G networks, it is crucial to optimize their performance. This paper explores the application of Artificial Intelligence (AI) and Machine Learning (ML) algorithms in the performance tuning of 5G networks, focusing on areas such as network resource allocation, traffic prediction, and real-time decision-making. By analyzing vast datasets generated by 5G infrastructures, AI and ML enable dynamic adjustments, thereby improving network efficiency, reducing latency, and enhancing overall user experience. The integration of these advanced technologies allows for self-optimizing networks that adapt in real time, minimizing human intervention and operational costs. This research highlights the key algorithms and techniques used for performance optimization, discusses the challenges of implementing AI in real-world 5G networks, and outlines the future directions for achieving fully autonomous network management. Ultimately, the study illustrates how AI and ML are pivotal in driving the future of telecommunications through intelligent 5G network tuning.

Machine learning for service classification in 5g networks

A challenge for future wireless communication networks is the satisfaction of the diverse requirements coming from heterogeneous services. In 5G networks, the coexistence of different services like Mobile Broadband (MBB), Massive Machine type Communications (MMC) and Mission Critical Communications (MCC) having various requirements in terms both of capacity and QoS will constitute a key prerequisite. Hence, one of the main is sues that should bead dressed by the 5G management system is the simultaneous provisioning of these services satisfying the corresponding requirements so as too ptimize the network in order to be resource and energy efficient. A first step towards this direction is to be able to identify each service type in order to prioritize the services and be able to allocate efficiently the networker sources.

Knowledge of QoS requirements per service flow could be provided by the higher layers as e.g. assumed in the HSPA and LTE, where sets of QoS parameters are available for RRM functionalities such as admission control and packet scheduling decisions. As an example from LTE, each data flow (bearer) is associated with a QoS profile consisting of the following downlink related parameters:

- Allocation retention priority (ARP)
- Guaranteed bitrate (GBR)
- QoS class identifier (QCI)

In particular, the QC Incudes parameters like the layer 2 packet delay budget and packet loss rate. However, for the cases where detailed QoS parameters are not made available from the higher layers, the use of novel service class if i cation techniques should be considered, in which the base stations monitor the traffic flows to extract more detailed service classification information and identify the service type providing this information input to packet scheduling algorithms, and other RRM functional it i es.

The support of fast and reliable traffic characterization is a necessary step in order to understand the network resource us age and to provide differentiated and high QoS/QoE through prioritization targeting in in creasing resource-usage and energy efficiency. In addition, the service classification process can interact with new services and procedures provided in 5G networks to support flexibility and adaptability to traffic variability.

In this subsection, the use of various machine learning mechanisms for the service classification problem is described and the performance of different algorithms is investigated. The considered classification methods reside in the area of statistical-based classification techniques and they are realized by exploiting several flow-level measurements (e.g. such as traffic volume, packet length, inter-packet arrival time and so forth) to characterize the traffic of different services. Then, to perform the actual classification, supervised machine learning techniques are applied to these measurements. It should be noted that in contrary to other methods of traffic classification, like payload-based classification, which need to analyze the packet payload or need to use deep packet inspection technologies, statistical-based classification techniques are usually very lightweight, as they do not access packet payload and can also lever age information from flow-level monitors.

State of the art of machine learning mechanisms for traffic classification

In the literature, there are a lot of studies that focus on application and service discrimination based on traffic classification learning techniques as presented in detail. Various machine learning mechanisms are usually employed belong ing to either unsupervised or supervised machine learning as illustrated in Figure 1. In the first case, clustering algorithms like K-Means, IDBSCAN and Auto class [6] are investigated. The objective of these mechanisms is to group flows that have similar patterns into a set of dis joint clusters. The major advantage of these schemes is that they do not require at raining phase like the supervised ones but they automatically discover the classes via the identification of specific patterns in the dataset. However, the resulting clusters do not certainly map 1: It o services a susu ally the number of clusters is greater than the number of service types and even in the case of 1:1 mapping, the clusters still need to be label ed in order to be map ped to the core responding services.

Regarding the supervised machine learning techniques, which is also the approach that is analyzed in this subsection, there are various classification schemes that have been proposed for the traffic classification problem like Naïve Bayes, Decision trees, Random for ests and others [7]. Authors in present the Bayesian classification techniques which use the Naïve Bayes approach. During the training phase, flow parameters are used to train the classifier and create a group of services. Then, when new flows arrive, they are subjected to probabilistic class assignment, by calculating their probabilities of class member ship and assigned to that class to which maximum probability is attained. In addition, statistical fingerprint-based classification techniques as presented in [9] classify traffic based on a set of pre-selected parameters (e.g. packet size, inter-arrival time). During the training phase, a dataset of flows from each service are used in order to analyze the data set and create the service finger print. This fingerprint is usually a PDFI (Probabilistic Density Function) vector used to identify the service. During the classification phase the algorithm checks the behavior of a flow against the available set of PDFI vectors. Also, Support Vector Machine (SVM) techniques, first proposed in [10], are binary supervised classificational girths which trans for man on linear classification problem in al in ear one, by mean sof what is called a “kerneltrick”. Further more, Artificial Neural Networks (NNs) cons is to f a collection of process in elements that are highly inter connected and transform a set of inputs to a set of desired outputs that is inspired by the way biological nervous systems works. In [11] author proposed NN, in which a multilayer perception classification network is used for assigning probabilities of lows. Aset of flow features are used as input to the first layer of network, while the output classify flow into a set of traffic classes by calculating the probability density function of class membership. Also, Decision Tree algorithms, which are mentioned also by [5], represent a completely orthogonal approach to the classification problem, using a tree structure to map the observation input to a classification outcome. In these supervised classification algorithms, the data set is learnt and modeled, there fore, whenever a new data item is given for classification, it will be classified accordingly learned from the previous dataset.

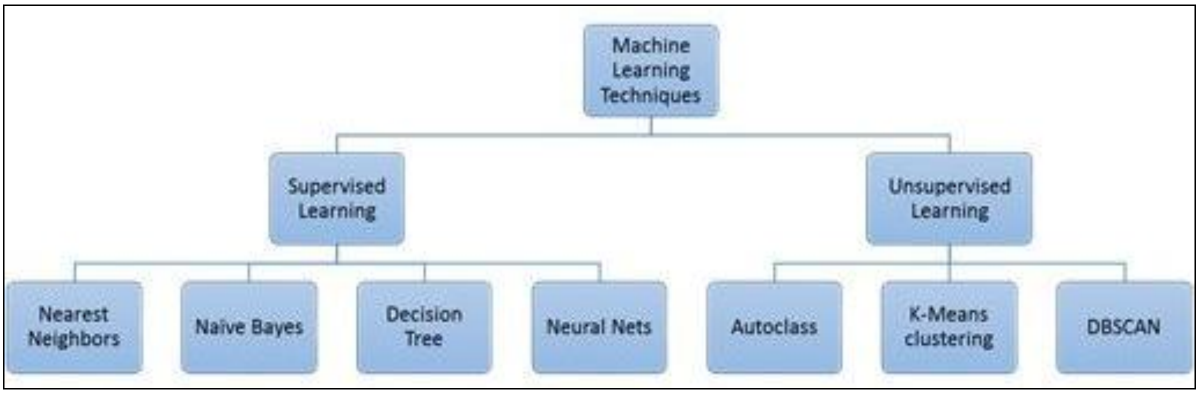

Fig 1: Overview of machine learning techniques

Classification approach & evaluation metrics

In this subsection, the problem of service classification is investigated employing a set of different classification mechanisms that belong to the supervised machine learning category. Before presenting the performance evaluation of each mechanism, the algorithm mic procedure that has been followed is described and each step is explained in detail. The algorithmic procedure that is followed for the classification mechanism. As can be seen, the first step refers to the collection of a number of traces from different services. For the considered simulation scenario, three service types are considered referring to MCC, MMC and MBB communication while other services (like broadcast/multicast services etc.) will also be considered in the future. The different traces of each service have been generated using specific traffic models. More specifically, the generation of different types of MCC/MMC traffic was following the traffic models presented in 802.16p [6], while for the generation of MBB traffic video streaming traffic (YouTube) which follows the traffic models presented in [7] is assumed.

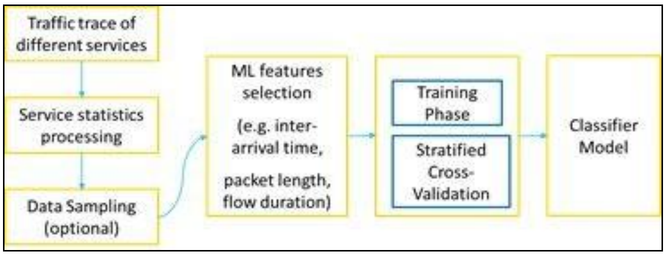

Fig 2: Proposed mechanism for the service classification process.

The second step refers to the statistical processing of these traces in order to separate them inflows. In particular, a flow is considered a series of packets transmissions that have the same source and destination and for which the inter-arrival time is below a specific threshold. After this processing, a number of features for each flow is generated including: inter-arrival time statistics (mean value, standard deviation), packet size statistics (total value, mean/max/min value, standard deviation) and other flow characteristics like total number of packets, source, destination as well as flow direction. Subsequently, some features engineering tasks are performed before proceeding to the main classification mechanism. These tasks include the selection of the most representative features, the transformation of categorical features into numerical values, the normalization of features' values and other tasks that guarantee a high data quality (e.g. replace missing values etc.). Then, the implementation of machine learning mechanism follows including two main phases: the training phase and the cross-validation phase. It should be noted that stratification is applied in this case in order to randomly sample the flows' data set in such a way that each service type is proper lyre presented in both training and testing datasets.

For the simulation scenario, a splitting of 70%-30% for training and testing sets has been considered. Obviously, for the training set, the label 'service type' of each flow is considered as known where as for the testing set, this label is considered as unknown and each flow is labeled using the classifier model. The outcome of the proposed mechanism is a classifier model that can be employed in unknown flows in order to recognize them and label them in an accurate way. to evaluate the performance of the classification mechanisms, various metrics have been defined and can be used in the train/test sets to select the most a de quate mechanism for the specific problem. To illustrate the relationship between the different evaluation metrics, a very useful tool that provides a holistic view of each algorithm's performance is the confusion matrix. The confusion matrix is actually a two dimension matrix, in which the horizontal axisre presents the predicted class (outcome of the algorithm) whereas the vertical axis represents the true class.

- FP: the percentage of other services' flows that are incorrectly classified as MMC service
- TP: the percentage of MMC service's flows that are correctly classified as MMC service
- FN: the percentage of MMC service's flows that are incorrectly classified as other services
- TN: the percentage of other services' flows that are correctly classified a sother services.

Classification Result Service	MMC service	Other Services
MMC Service	TP	FN
Other Services	FP	TN

Fig 3: Confusion Matrix of the service classification problem.

Some of the most common evaluation metrics used for classification problems are the accuracy metric, the precision, the recall and the F1-score. More specifically:

- Accuracy is defined as the percentage of correct predictions to the total number of prediction sand is given by
- Precision is defined as the percentage of the instances that were correctly predicted as belong in gin a class among all the in stances that were classified as be longing in this class and is given by

$$\frac{(TP+TN)}{(TP+FP+TN+FN)}$$

- Recall is defined as the percentage of the instances of a specific class that were correctly classified as belonging to this class and is given by

$$\frac{TP}{(TP+FN)}$$

- F1 Score is defined as the harmonic mean of the precision and recal land is given by

$$\frac{2 \times \text{Precision} \times \text{Recall}}{(\text{Precision} + \text{Recall})}$$

- To be able to choose the best mechanism for a classification problem, the investigation of a single metric, like accuracy, is not always enough as the misclassification of a specific class instances maybe more important than the correct classification of others. For this reason, other evaluation metrics have also to be applied to make the most appropriate choice depending on the problem's characteristics.

Evaluation performance of classification mechanisms

In the considered simulation scenario, the performance of a set of different machine learning mechanisms has been investigated including base classifiers such as Naïve Bayes classifier, Support Vector Machines, Tree Classifier, K Nearest Neighbor Classifier, Logistic Regression as well as ensemble based classifiers like Random Forest Classifier. The goal of ensemble methods is to combine the predictions of several base estimators built with a given learning algorithm in order to improve generalizability and robustness over a single classifier. Usually, two families of ensemble methods are distinguished: the averaging methods (e.g. Random Forests) [8] in which several classifiers are developed independently and then the average of their predictions is used and the boosting methods (e.g. Ada Boost-Adaptive Boosting) where base classifiers are built sequentially and one tries to reduce the bias of the combined estimator. To be able to compare the various machine learning mechanisms, in the following table, the accuracy metric of each algorithm is presented, where a Dump classifier that classifies all the flows as type (MMC service) is also considered resulting in 0.512 accuracy. From this table, it can be seen that Decision tree and the Random Forest algorithms lead to the highest accuracy values, out performing the other machine learning algorithms.

Table 1: Accuracy score for each classification mechanism.

Classification Mechanism	Accuracy
Naive Bayes	0.808
Support Vector Machine	0.662
Decision Tree	0.976
K Nearest Neighbor Classifier	0.952
Logistic Regression	0.685
Random Forest Classifier	0.988

However, to provide a more complete view of each classifier's performance, the corresponding confusion matrices are illustrated. The horizontal axis of this matrix re presents the predicted class whereas the vertical axis represents the true class. It should be noted that Class 0, Class 1 and Class 2 refer to MMC, MCC and MBB service types, correspondingly. In the considered scenario, considering that it is desired to eliminate the possibility that a MCC service is misclassified as another service type, the optimal model should have high values of recall whereas high accuracy values for the case of MMC and MBB services are required. The results of confusion matrix show that the Decision Tree and the Random Forest algorithms result in extremely good results as they misclassify only a few flows, resulting also in high values of recall. There fore, these two classification mechanisms can be selected chosen for further consideration for the problem of service classification.

Fig 4: Confusion matrices of different classifiers

Class\ Metrics	Precision	Recall	F1-score	Class\ Metrics	Precision	Recall	F1-score
MMC	0.99	1.00	0.99	MMC	0.99	0.97	0.98
MCC	0.99	0.97	0.98	MCC	0.94	0.97	0.96
MBB	0.99	0.99	0.99	MBB	0.98	0.99	0.99
Avg/total	0.99	0.99	0.99	Avg/total	0.98	0.98	0.98

Fig 5: Evaluation metrics for selected classification mechanisms.

Conclusion

5G is the next frontier of innovation for entire mobile industry. Consequently the three major objectives for 5G, are support of massive capacity and massive connectivity; support for an increasingly diverse set of services, applications and users; and in addition flexible and efficient use of all available non-contiguous spectrum for widely different network deployment scenarios. Framed in this context, this chapter elaborated on the status & challenges in hardware/software development and in 5G wireless communications by focusing on physical layer, MAC and RRM. Also the benefits of machine learning in 5G network management were discussed. By taking into account the diversity of infrastructure, radio resources and services that will be available in 5G, an adaptive network solution frame work will become a necessity. Break through developments in several RAN technologies will be required for realizing novel, 5G solutions. Such technologies include among others multi places and advanced wave form technologies combined with coding and modulation algorithms, massive access protocols, massive MIMO and virtualized and cloud-based radio access infrastructure.

References

1. Singh S. The role of artificial intelligence in 5G network management: A comprehensive study. International Journal of Research Publication and Reviews; c2023. p. 2862-2866. doi:10.55248/gengpi.4.1123.113203.
2. Barakabitze A, Hines A. 5G network management for big data streaming using machine learning. In: Wiley Handbook of Science and Technology for the Internet of Things. 2022. doi:10.1002/9781119800828.ch2.
3. Hernández-Chulde C, Cervelló-Pastor C. Intelligent optimization and machine learning for 5G network

control and management. In: Lecture Notes in Computer Science. Springer; c2019. p. 33. doi:10.1007/978-3-030-24299-2_33.

4. Kaur J, Khan MA, Iftikhar M, Imran M, Haq Q. Machine learning techniques for 5G and beyond. IEEE Access. 2021;PP:1-1. doi:10.1109/ACCESS.2021.3051557.
5. Khan I, Joshi A, Antara F, Singh D, Goel O, Jain S. Performance tuning of 5G networks using AI and machine learning algorithms. International Journal for Research Publication and Seminar. 2020;11:406-423. doi:10.36676/jrps.v11.i4.1589.
6. Elliott J, Sharma S. Dynamic end-to-end network slicing unlocks 5G possibilities. Available from: https://insight.nokia.com/dynamic-end-end-network-slicing-unlocks-5g-possibilities. Accessed Nov 2016.
7. Bogale T, Le I. Massive MIMO and millimeter wave for 5G wireless HetNet: potential benefits and challenges. IEEE Vehicular Technology Magazine. 2016;11(1):64-75.
8. Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-III. IEEE Transactions on Evolutionary Computation. 2002;6(2):182-197.
9. Saur S, Weber A, Schreiber G. Radio access protocols and preamble design for machine-type communications in 5G. In: Signals, Systems, and Computers, 49th Asilomar Conference. 2015. p. 3-7.
10. 3GPP. Further advancements for E-UTRA physical layer aspects. 3GPP Technical Report 36.814. March 2010.
11. Namdev N, Agrawal S, Silkari S. Recent advancement in machine learning based internet traffic classification. Procedia Computer Science. 2015;60:784-791.
12. Valentini S, *et al*. Reviewing traffic classification. In: Data Traffic Monitoring and Analysis. Springer; c2013. p. 123-147.
13. Nguyen TT, Armitage G. A survey of techniques for internet traffic classification using machine learning. IEEE Communications Surveys & Tutorials. 2008;10(4):56-76.
14. Aafa JS, *et al*. A survey on network traffic classification techniques. International Journal of Engineering Research and Technology. 2014;3(3).
15. ITU. Guidelines for evaluation of radio interface technologies for IMT-2020. Document 5/57. ITU; c2013.