INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
Online ISSN 1841-9844, ISSN-L 1841-9836, Volume: 16, Issue: 6, Month: December, Year: 2021
Article Number: 4583, https://doi.org/10.15837/ijccc.2021.6.4583

CCC Publications

Improving the Performance of Heterogeneous Network Systems in Machine Learning-based 5G Mobile Communication System

Y. H. Kim, D. Y. Lee, S. H. Bae, T. Y. Kim

Yoon Hwan Kim, Sang Hyun Bae
Department of computer Science & Statistics
Chosun University, Rep. of KOREA
309 Pilmun- Daero, Dong-gu, Gwangju 61452, Rep. of KOREA
kyh0301@ccn.ac.kr, shbae@chosun.ac.kr

Dae Young Lee
Civil Communication Planning Division
GwangJu Metropolitan City, Rep. of KOREA
111 Naebang-ro, Seo-gu, Gwangju, 61945, Rep. of KOREA
ldy0722@korea.kr

Tae Yeun Kim*
National Program of Excellence in Software center
Chosun University, Rep. of KOREA
309 Pilmun- Daero, Dong-gu, Gwangju 61452, Rep. of KOREA
*Corresponding author: tykim@chosun.ac.kr

Abstract

Mobile traffic, which has increased significantly with the emergence of Fourth generation long-term evolution (4G-LTE) communications and advances in video streaming services, is still currently increasing at an incredible pace. Fifth-generation (5G) mobile communication systems, which were developed to deal with such a drastic increase in mobile traffic, aim to achieve ultra-high-speed data transmission, low latency, and the accommodation of many more connected devices compared to 4G-LTE systems. 5G communication uses high-frequency bandwidth to implement these features, which leads to an inevitable drawback of a high path loss. In order to overcome this disadvantage, small cell technology was developed, and is defined as small, low-power base stations that can extend the network coverage and solve the shadow area problem. Although small cell technology has these advantages, different problems, such as the effects of interference due to the deployment of a large number of small cells and the differences in devices accessing the network, need to be solved. To do so, it is necessary to develop an algorithm for a service method. However, general algorithms have difficulties in responding to the diverse environment of mobile communication systems, such as sudden increase in traffic in certain areas or sudden changes in the mobile population, and machine learning technology has been applied to solve this problem. This study employs a machine learning algorithm to determine small cell connections. In addition, a 5G macro system, the application of small cells, and the application of machine learning algorithms are compared to determine the performance improvement in the machine learning algorithm. Moreover, Support Vector Machine

(SVM), Logistic Regression and Decision Tree algorithm are employed to show a training method that uses basic training data and a small cell on-off method, and the performance enhancement is verified based on this method.

Keywords: machine learning, logistic regression, support vector machine, decision tree, 5G, small cell.

1 Introduction

Since the emergence of fourth generation (4G) LTE communication systems, mobile traffic has significantly increased owing to the video content market, such as YouTube.

Fifth-generation (5G) mobile communication technology, which has emerged due to the sudden increase in mobile traffic, has evolved from existing speed-centered technologies and aims to satisfy the high speed, low latency, high energy efficiency, and connection of a large number of devices. It is expected that mobile traffic will increase even more with the commercialization of 5G communication.

Up to November 2020, 122 mobile network operators in 49 countries had launched commercial 5G services. Among them, 115 mobile network operators provide 5G mobile services, and 40 mobile network operators provide 5G fixed wireless access (FWA) or broadband services [1].

The 5G mobile communication systems that have been developed utilize the 3.5-GHz frequency band, which is higher than that of 4G mobile communication systems, and it ultimately aims to achieve maximum transmission capability using millimeter-wave technology. However, because 5G mobile communication uses high frequency, it is susceptible to interference owing to moisture in the air or windows, which is a limitation of 5G mobile communication technologies [2].

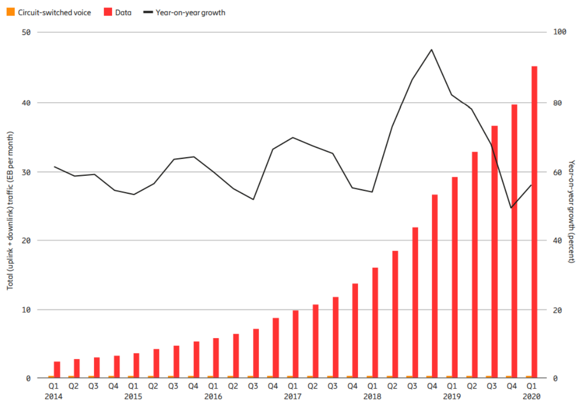

Figure 1: Global Mobile Network Data Traffic and Year-on-year Growth [3]

To overcome this drawback, the rollout of a large number of 5G base stations could be deployed, but this approach would have cost implications, and communication inside buildings would also be limited.

Small cell technology, which is a technology that can enhance the system performance by extending the coverage and employing frequency reuse, even in existing LTE systems, is a potential solution that can improve this problem in 5G mobile communications. The use of a small cell is a method that uses the band licensed to existing cellular systems. This method presumes an ultra-dense network (UDN), which deploys a large number of small cells. Hence, there is a lot of interference, which needs to be managed. If not, the performance of the entire system will be degraded. In this study, a machine learning-based system was constructed for the basic small cell application, to improve the small cell application algorithm, and to achieve small cell connection and interference management. The results confirmed that there was a large performance improvement over existing 5G mobile communication using small cell technology and machine learning technology [4].

This paper is organized as follows. Section 2 examines related technologies, and Section 3 describes the system model and the proposed method. Section 4 analyzes the simulation results, and Section 5 concludes the paper.

2 Related Technology

2.1 Fifth generation mobile communication

First generation (1G) mobile communication refers to a technology that enabled voice call services with a transmission speed of 9.6 14.4 kbps, while second generation (2G) mobile communication is a code-division multiple access (CDMA)-based voice and text service with a transmission speed of 14.4 64 kbps. With the development of third generation (3G) mobile communication, the transmission speed was improved to 2.4 Mbps, making video calls possible. Then, 4G mobile communication was released and provided speeds of up to 1 Gbps, and which enabled the development of the mobile streaming market. In the fifth generation (5G) of mobile communication, a maximum of 50 Gbps is supported. Hence, video content services dominate the mobile traffic market, as shown in Figure 2, and an exponential increase in mobile traffic is anticipated [3].

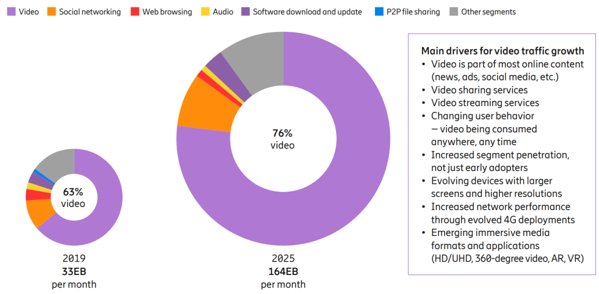

Figure 2: Mobile Traffic by Application Category Per Month (Percent) [3]

The 5G specifications were developed based on LTE-Advanced (LTE-A) Pro (4.5G) technology, and the 5G standard was named 5G Phase 1 in 3GPP Release-15. It is also called New Radio (NR) because it utilizes new frequencies in the 3.5-GHz band and the 28-GHz band rather than the existing frequencies owing to the lack of available frequencies in the existing LTE band.

5G mobile communication is defined by three service categories: enhanced Mobile Broadband (eMBB), Ultra-Reliable & Low Latency Communications (URLLC), and massive Machine-Type Communication (mMTC) [5].

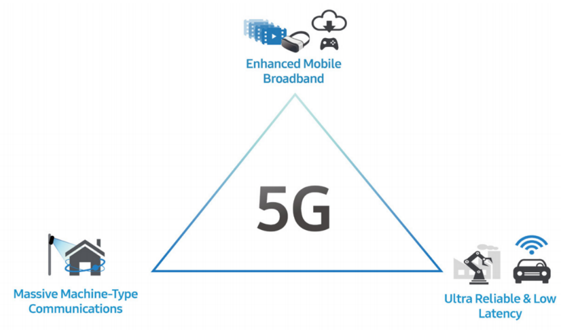

Figure 3: 5G Vision [5]

In 5G, eMBB is a technology that enables the transmission of a large amount of data using a large bandwidth. It uses a bandwidth of 400 MHz at 28 GHz, which is defined as a millimeter wave, to support speeds that are eight times faster than that of existing 4G LTE services. The latency of 4G LTE is 10 ms, and URLLC improves this latency by a factor of more than 10 times to 1 ms in order to achieve high reliability and ultra-low latency. URLLC is a technology that enables remote control of things, such as remote medical services. Moreover, mMTC allows Internet of Things (IoT) devices or numerous devices to connect. The goal of mMTC is to support one million connections per square kilometer [5].

2.2 Path loss in high-frequency band

The ultimate objective of 5G mobile communication is to enable ultra-high-speed, ultra-low latency services using high frequencies. However, the use of high frequencies results in a very large path loss. As shown in the following equation for path loss, the power of the received signal Prx generally experiences a path loss corresponding to the square of the frequency with respect to the power of the transmitted signal Ptx [6].

$$P_{rx} = P_{tx}G_{tx}G_{rx}\left(\frac{\lambda}{4\pi R}\right)^2 = P_{tx}G_{tx}G_{rx}\left(\frac{c^2}{4\pi * f^2}\right)\left(\frac{1}{4\pi R^2}\right) \quad (1)$$

Gtx is the transmitter antenna gain, Grx is the receiver antenna gain, and R is the distance between the transmitter and the receiver. The path loss values for 2.8 GHz and 28 GHz are -41.4 dB and -61.4 dB, respectively. From this comparison, it can be seen that the path loss value is huge when using a high-frequency band [7].

In 5G systems, the maximum bandwidth is 50/100/200 MHz at 3 GHz and 200/400 MHz at 28 GHz. In 4G-LTE, the frequency division duplex (FDD) method is utilized, so the upload and download bands are used separately. As a result, interference does not occur, but it does have the drawback of having poor frequency efficiency. However, the time division duplex (TDD) method is utilized in 5G systems, so the upload and download bands are not divided. Instead, a part of the entire frequency bandwidth can be divided and used according to the time slot; hence, it can accommodate a large number of connected devices [8].

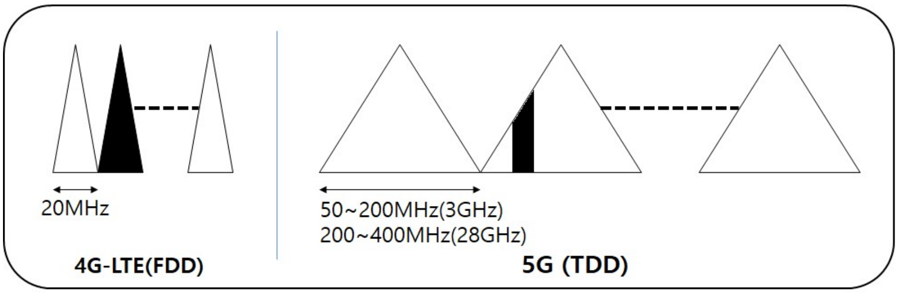

Figure 4: Comparison between FDD and TDD [8]

2.3 Small Cell

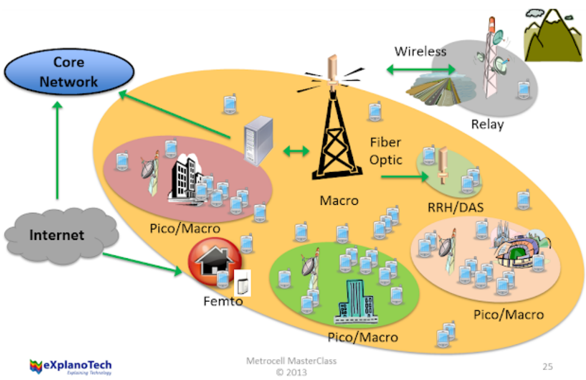

Figure 5: Heterogeneous Networks [11]

Small cells are low-power base station equipment and have the following advantages. Compared to macro base stations, small cells are less expensive to build and are smaller. Hence, it is easy to build a small cell within a particular space in a short time. In addition, because small cells are low-power base stations, they have a low maintenance cost. Small cells are classified as femtocells (21 dBm), picocells (24 dBm), and microcells (38 dBm) according to the output power [9].

A small cell is defined as a small cell technology in the 3GPP Release.10 and Release.11, and it is a technology that is expected to reduce shadow areas or increase transmission capacity by setting up multiple access points (APs) within a cellular system, and it can improve system performance by employing frequency reuse in the cellular system. In 5G systems, the large path loss encountered owing to the use of high frequencies results in degraded communications quality. This phenomenon can be improved by applying small cell technology [10].

2.4 Machine learning

Machine learning is a technology in which a computer learns the patterns or rules of data on its own using a machine-learning algorithm without programming, and predicts the results for new data. The learning algorithm for machine learning utilizes artificial neural networks (ANNs), and is typically classified into supervised learning and unsupervised learning. In this section, we are considering in the paper supervised learning.

2.4.1 Supervised learning

Supervised learning is further classified as regression and classification. Supervised learning learns by the set of input data X and label set Y to be learned, and supervised learning has two steps: training and prediction. Common classification models include k-nearest neighbor (kNN), support vector machine (SVM), and decision tree. In addition, regression models that predict results for a given input are important. In regression models, logistic regression predicts categorical outcome values. In unsupervised learning, data labels are not provided for training [12].

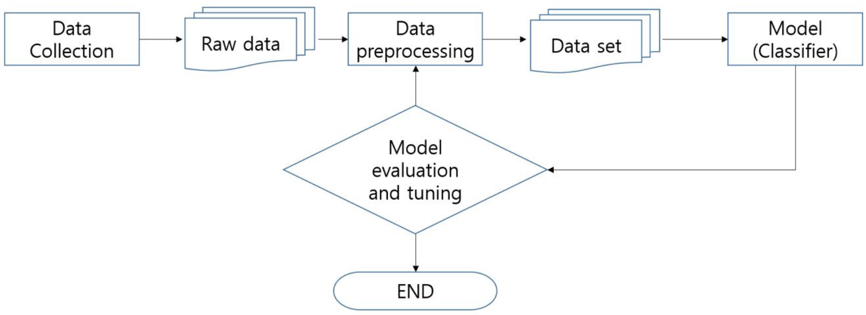

Figure 6: Machine Learning Classification Model Stage [13]

Figure 6 is a diagram showing the machine learning classification model stage. In the step of machine learning classification learning, data is collected first, and then the collected raw data is preprocessed. If you do not go through this data preprocessing step, noise or unnecessary data will be introduced, which will affect the performance of the machine learning model. Data preprocessing goes through data reduction to select or delete overlapping data and conversion of raw data into a specified format as needed by the model, and data integration and data organization. Then, based on the created dataset, learning is conducted with a machine learning model, and evaluation of the learned model is conducted [13].

2.4.2 Machine learning performance measurement

Three methods were used to measure and analyze the machine learning performance, namely. accuracy, recall, and F1-score. Formulas for each performance analysis method are summarized as follows by referring to Table 1.

- TP (True Positive): A value that predicts an actual positive value as a positive value.
- TN (True Negative): A value that predicts an actual negative value as a negative value.

- FP (False Positive): A value that predicts an actual negative value as a positive value.
- FN (False Negative): A value that predicts an actual positive value as a negative value.

Accuracy is defined by Equation (2), and is the ratio of the number of correct answers to the total number of predictions. If the proportion of negative values is high in the actual data of the model, it does not produce an accurate accuracy value. In this case, the recall of Equation (4) is used. The recall is the ratio of the number of values that have been predicted as true to the number of values for which the correct answer is true.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

Table 1: Confusion matrix

		Predicted	
		Positive	Negative
Actual	Positive	TP (True positive)	FN (False negative)
	Negative	FP (False positive)	TN (True negative)

2.4.3 Logistic regression algorithm

If the problem to be solved is a classification problem, it is solved by finding the monotonic differentiable function to connect the actual labels of the classification problem with the predicted values of the linear regression model. Moreover, the model for classification possibility is created directly, and assumptions about the prior data distribution are not needed [14].

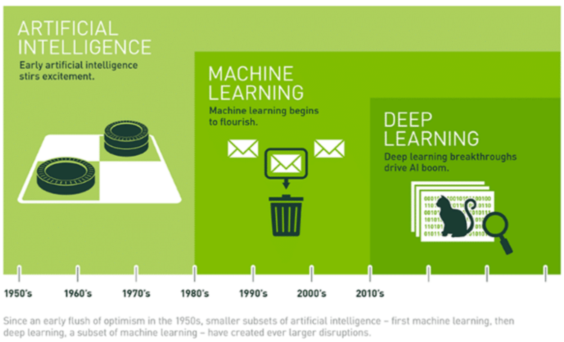

Figure 7: Artificial Intelligence [15]

2.4.4 Decision tree algorithm

A decision tree is similar to the human thinking and decision-making process. A decision tree is generally composed of one root node, several internal nodes, and leaf nodes under the internal nodes. The sample set of each node is a structure that is classified into lower nodes after testing, and it has outstanding generalization performance [14].

2.4.5 Support vector machine algorithm

The SVM is a model that determines the decision boundary. When new unclassified data appear, SVM solves the classification problem by determining the side of the boundary to which the new data belongs. In SVM, the decision boundary is defined through support vectors, so unused data points can be managed efficiently [14].

3 System model

3.1 System parameter

The main base station exists in the center, and small cells are placed randomly within the cell. The radius of the base station is 1 km, and an environment where the base station is subject to interference from six neighboring base stations is assumed. The small cell communication outage is less than 10% up to 10 m, based on the low output of the small cell. However, the outage increases drastically when the distance exceeds 20 m. As a result, connections are made within 10 m by default for small cell communications.

Table 2: System parameter

Parameter	Value
Cell radius	1km
eNB Power	20W
Small Cell Power	10W
Device Power	200mW
UE Device	10~100
Distance between SBS-SBSUE	10m
Distance between D2D	20m
Distribution Device	Random
Simulation count	500
Pathloss	eNB : 15.3+37.6logD
	SBS : 38.46+20logD
	D2D : 148+40logD
Noise density	-174dBm/Hz
eNB	25m
Simulation Machine	AMD-3970X
	128GB RAM
	Nvidia RTX4000
Simulation Program	Matlab 2018a

3.2 Small cell machine learning algorithm

Table 3: 5G Small cell machine learning algorithm

	Algorithm : 5G Small Cell Machine Learning Algorithm
1	User Number : N
2	Input Data : SBS1, MBS1, User1, SINR, Capacity
3	Data Generate, Labeling
4	Begin (Data, Label, Data Set)
5	Creating Model
6	Train (Model, Data, Label)
7	While All Users Do
8	Result = Predict
9	Small Cell Connection (Result)
10	End While
11	End

The general placement of small cells or the application of algorithms could be a factor that degrades the system performance owing to variables in the actual environment. For example, when determining the use of the macro cells and small cells using the signal-to-interference-noise ratio (SINR) or capacity, connections are made to the base station, where the optimal value is calculated by the algorithm. However, the system performance could be degraded because a large number of mobile devices are connected to a particular base station. To solve such a problem, a method was developed so that the system could learn by itself to find the optimal method of determining small cell connections using a machine-learning algorithm. In order to use the machine-learning algorithm, 50,000 sets of small cell communication simulation data were used and transformed into a format appropriate for machine learning. Next, the data were divided into 40,000 data sets for training and 10,000 data sets for testing.

In this study, a supervised learning method was used to create a model that can make selections autonomously by training the model with the correct answer after labeling the communication method at the corresponding location.

For the machine-learning algorithm, SVM, logistic, and decision tree algorithms were used as the classification method of the supervised learning to determine the small cell connection method. Further, a dual training method was applied to perform the simulation. The basic dataset of the small-cell machine-learning algorithm consists of a macro base station location, small cell base station location, user location, SINR, and capacity values. The labeling task is performed in a way that selects the base station with high throughput at the user's location by calculating the SINR and capacity.

3.3 Support vector machine algorithm classification prediction

The sequential minimal optimization (SMO) method was used to apply the SVM algorithm to classify the communication connection method of users. In practice, it is difficult to have a hyperplane that can completely separate the training data when running the algorithm. To alleviate this problem, the soft margin SVM was used to determine a decision boundary that allows some misclassification.

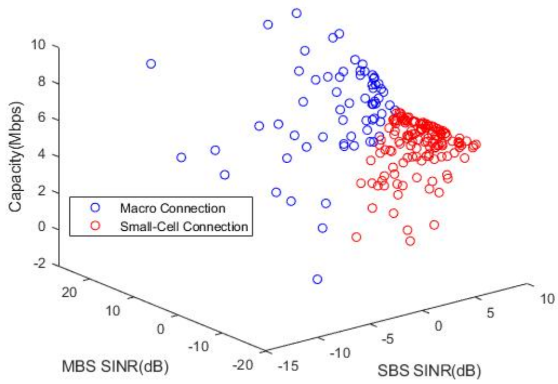

Figure 8: Small Cell SVM Algorithm Training Model Classification

3.4 Logistic regression algorithm classification prediction

For the application of the logistic algorithm, the parameter value that minimizes the value of the cost function was applied using the gradient descent method when the small-cell machine-learning algorithm is run. When running the logistic algorithm, the values for small cells and macro cells are converted into binary numbers through labeling in the data preparation phase, and are delivered as True or False values.

3.5 Decision tree algorithm classification prediction

The construction of the decision tree algorithm for estimating the communication connection method was adjusted to set the training data and test data and the maximum number of splits. The use of 100 splits produced a result of 92.8%, which was the highest value. 20 splits produced 91.8%, 50 splits produced 92.1%, and 200 splits produced 92.6%. When the decision tree has split a maximum of 1000 times, it produced a result of 90.1%; the accuracy increased to a certain extent before decreasing. The split criteria include the Gini impurity and entropy impurity. Because the difference between the result values of the two methods was not large, the Gini impurity was used. Figure 9 shows a schematic diagram of the tree shape resulting from the decision tree training. The yellow color indicates macro communication, and the green color indicates small cell communication. The Gini value is close to 0 in the decision tree, indicating the proper selection of the communication connection method.

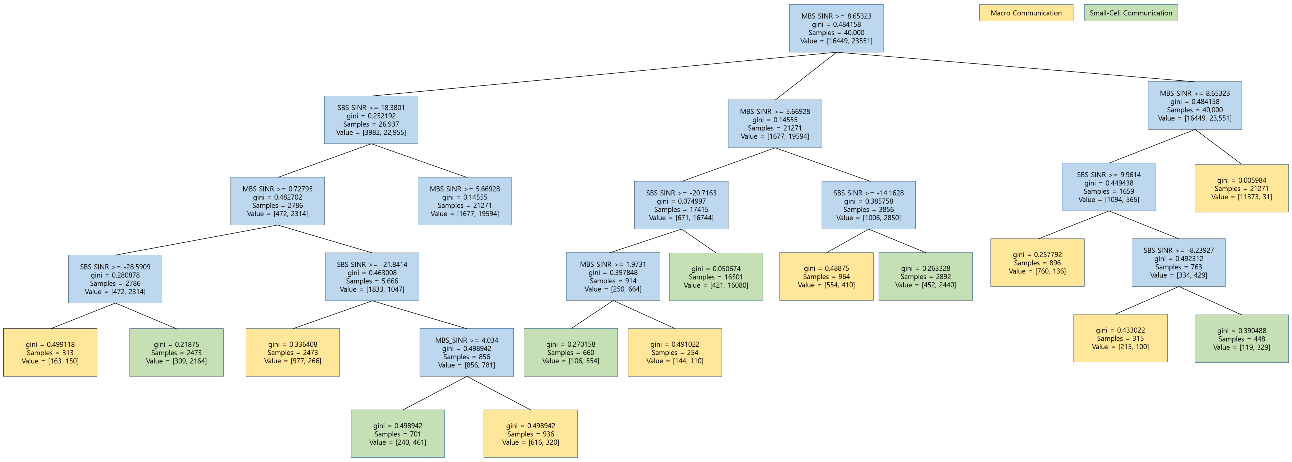

Figure 9: Decision Tree Training Model

4 Simulation results

4.1 5G Small cell performance comparison

Figure 10 shows a graph comparing the case involving the use of only the macro cells for communication in 5G communication and the case involving the application of small cell technology. When small cells are applied to the macro system, the performance is better than the performance of the existing macro system.

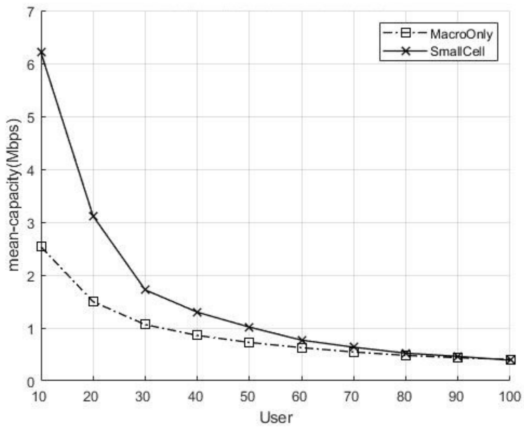

Figure 10: Capacity comparison when 5G macro communication and small cell are applied

4.2 5G Small cell application simulation result

Figure 11 shows the locations of the small cells, users, and activated small cells when small cells are applied to the 5G macro communication system.

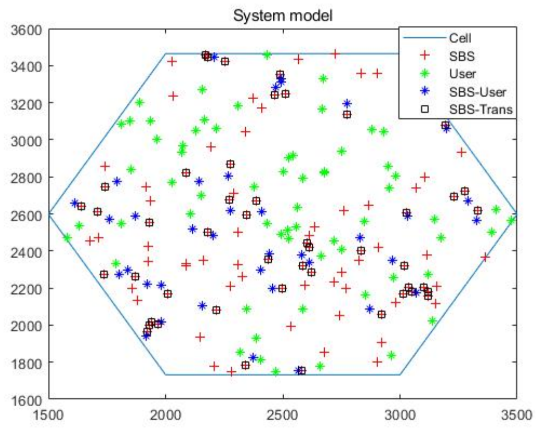

Figure 11: 5G Small cell system

4.2.1 Accuracy analysis using 5G machine learning algorithms

Figure 12 shows a graph of the accuracy of each machine-learning algorithm for the mobile communication system employed in this study when small cells have been applied. The result values were calculated by comparing the number of values for which the data labels and the prediction of the trained model match. In this paper, the data were calculated by comparing the data labels with the values predicted by the model for the machine-learning algorithm, which determines the macro cell communications and small cell communications. The proposed SVM model generated the highest prediction value of 0.9347, the logistic model produced a low prediction value of 0.9003, and the decision tree model produced a prediction value of 0.9282. As mentioned earlier in the application of the small cell machine-learning algorithm, the analysis results of the accuracy obtained by applying the dual training method were 0.9820 for the SVM algorithm, 0.9384 for the logistic algorithm, and 0.9632 for the decision tree algorithm. The results show large performance improvements for all three algorithms, and the SVM algorithm's performance improved the most.

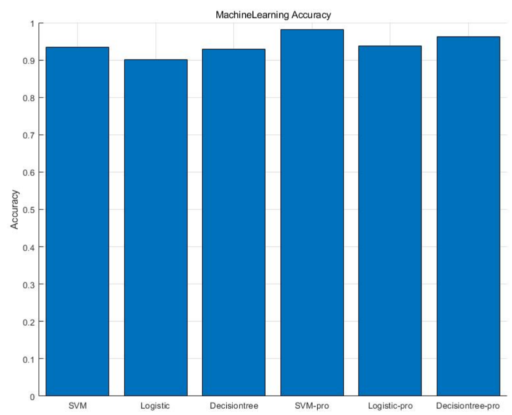

Figure 12: Machine Learning Accuracy (Small cell System)

4.2.2 Analysis of recall for small cell machine learning algorithms

Figure 13 shows the graph of recall according to the application of the small cell machine learning algorithms. The recall analysis shows the comparison of the predicted values with the actual values for the small cell connections. The logistic algorithm had the lowest recall of 0.9499, and the decision tree algorithm had the highest recall of 9.845; the SVM algorithm had a recall of 9.518. After performing the dual training, the recall for the SVM algorithm was 0.9940, for the logistic algorithm it was

0.9747, and for the decision tree algorithm it was 0.9901. As shown in the graph, there was an overall performance improvement for recall with the dual training.

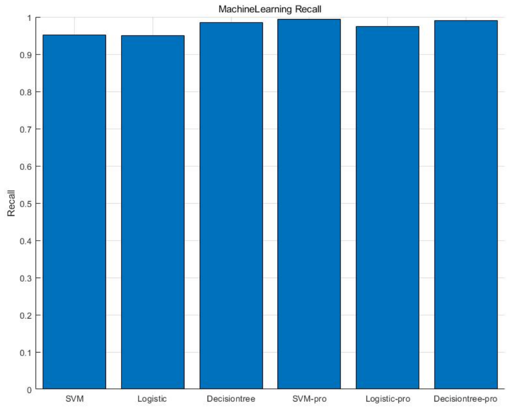

Figure 13: Machine Learning Recall (Small Cell System)

4.2.3 Analysis of precision for small cell machine learning algorithms

Figure 14 shows the precision results of the small cell machine learning algorithms. The SVM algorithm has the highest precision of 9.786. The logistic algorithm and decision tree algorithm have similar precision values of 9.404 and 9.405, respectively. After performing the dual training, the result was 0.9866 for the SVM algorithm, 0.9404 for the logistic algorithm, and 0.9709 for the decision tree algorithm. Prior to performing the dual training, the SVM had the highest recall, and the logistic and decision tree algorithms had similar recall performance values. However, the decision tree algorithm exhibits a better performance than the logistic algorithm after the dual training.

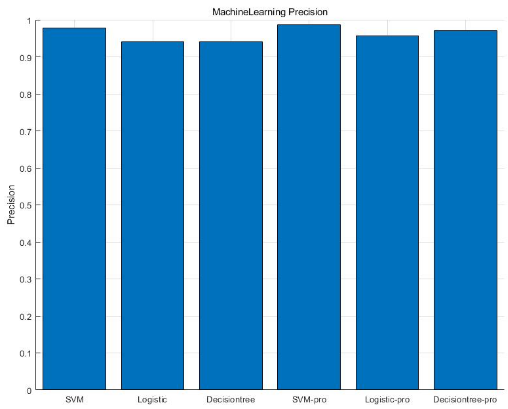

Figure 14: Machine Learning Precision (Small Cell System)

4.2.4 Analysis of F1-score for small cell machine learning algorithms

Figures 15 shows the performance evaluation of the small cell machine learning algorithms that were obtained using the F1-score. The SVM algorithm has the highest value of 9.650, the logistic algorithm has the lowest value of 9.450, and the decision tree algorithm has a value of 9.619, which is similar to the performance of the SVM algorithm. The results obtained by applying the dual training are 0.9903 for the SVM algorithm, 0.9450 for the logistic algorithm, and 0.9804 for the decision tree algorithm. The results show the improvement of each performance.

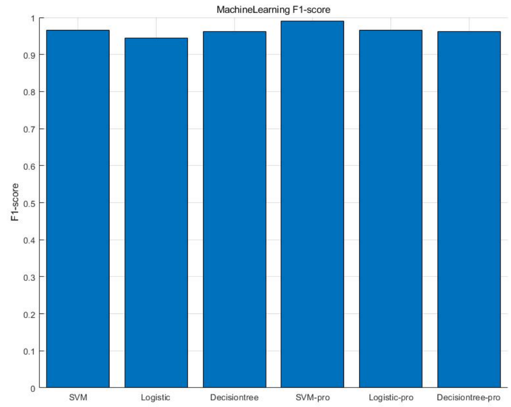

Figure 15: Machine Learning F1-score (Small Cell System)

4.3 Comprehensive analysis of learning speed and performance for the small cell machine learning algorithms

Table 4: Small cell machine learning algorithm learning speed

Algorithm	Learning Speed (sec)
SVM	45.576
Logistic Regression	23.678
Decision Tree	6.472
SVM (Dual training)	74.982
Logistic Regression (Dual training)	38.930
Decision Tree (Dual training)	11.487

The four performance analysis methods (accuracy, recall, precision, and F1-score) were described earlier for the small cell machine learning algorithms. In this section, the learning speed and the overall performance of the small cell machine learning algorithms are analyzed. For each performance method of the small cell machine learning algorithms besides the recall method, the SVM algorithm showed the best performance, followed by the decision tree algorithm. The decision tree algorithm showed a high performance in the recall. For the logistic algorithm, the performance was the lowest. When the dual training was applied, the overall performance was improved, which was confirmed in the above sections. In the precision analysis, a greater performance improvement was observed in the decision tree algorithm compared with the logistic algorithm when the dual training was applied. When implementing machine learning algorithms, the learning speed of machine learning can play a very important role in an environment that changes in multiple ways. Table 4 shows the learning speed of the small cell machine learning algorithms. The SVM algorithm is superior in each performance analysis method, but it is somewhat slow with respect to the learning speed. However, the decision tree algorithm has the fastest learning speed, and it has a better performance than the SVM algorithm with respect to the recall method. In addition, the decision tree algorithm shows performances that are similar to that of the SVM algorithm in terms of the accuracy and F1-score methods. Considering these results, the decision tree algorithm is considered to be superior in terms of its overall performance, including the learning speed. Moreover, the overall learning speed increases when dual training is applied. The application of small cell algorithms should therefore be configured based on the indicators that are determined according to each configuration environment.

4.3.1 Results of SVM algorithm simulation

Figure 16 shows the capacity when the decision tree algorithm is applied to the small cells. This system was compared with a general small cell system in the same way as the algorithms described earlier, and the performance improved from 1.20 times to 2.17 times.

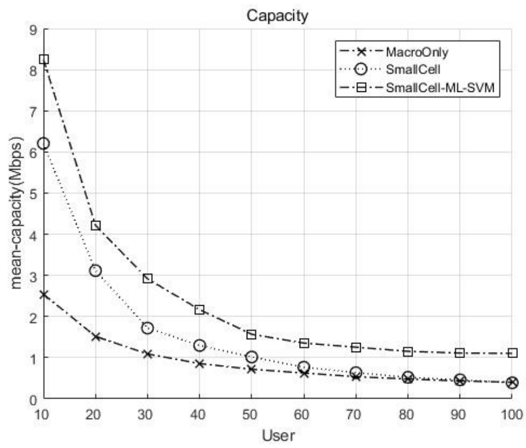

Figure 16: 5G Small Cell Machine Learning Algorithm Average System Total (SVM)

4.3.2 Results of logistic regression simulation

Figure 17 shows the results obtained by applying the logistic regression algorithm among the small cell machine learning algorithms. This system was compared with a general small cell system in the same way as the algorithms described earlier. The result showed a performance improvement of 1.06 times to 1.68 times. The performances were similar when there were 10 to 20 users, but a difference can be observed from 30 users.

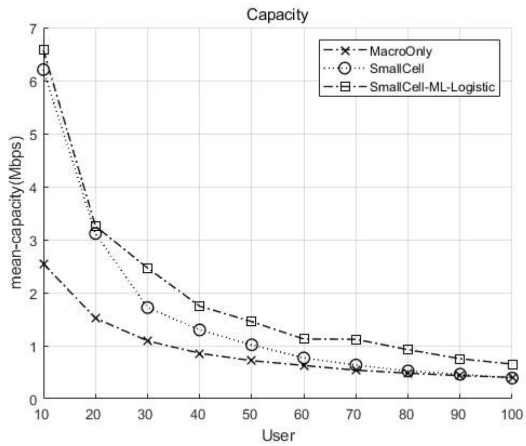

Figure 17: 5G Small Cell Machine Learning Algorithm Average System Total (Logistic Regression)

4.3.3 Results of decision tree simulation

Figure 18 shows the capacity when the decision tree algorithm is applied to the small cells. This system was compared with a general small cell system in the same way as the algorithms described earlier, and the performance improved from 1.20 times to 2.17 times.

4.3.4 Dual training simulation results for SVM algorithm

Figure 19 shows a graph comparing the capacities of the SVM algorithm among the small cell machine learning algorithms described earlier, the proposed SVM algorithm, and a general small cell system. When the proposed SVM algorithm is used, the performance is improved by 1.34 times to 3.10 times. Compared to the SVM algorithm, the performance improved by 1.01 times to 1.31 times.

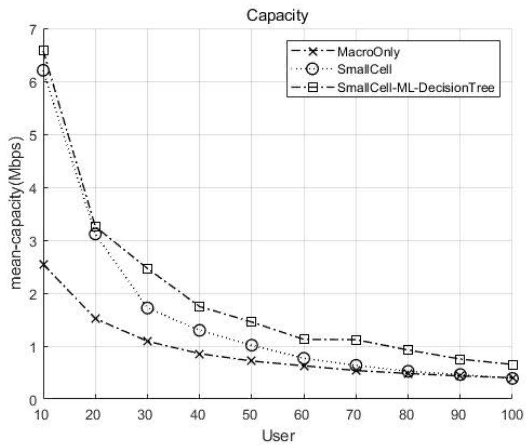

Figure 18: 5G Small Cell Machine Learning Algorithm Average System Total (Decision Tree)

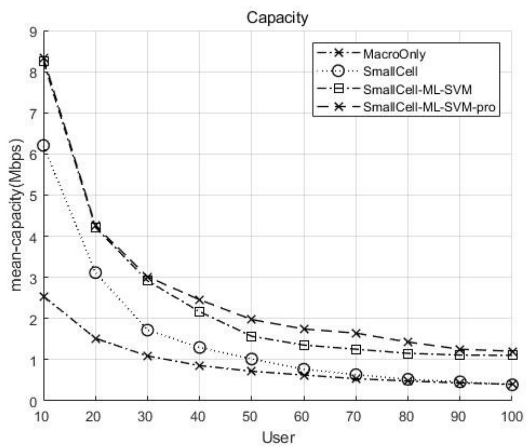

Figure 19: 5G Small Cell Machine Learning Algorithm Average System Total (SVM Dual Training)

4.3.5 Dual training simulation results for logistic regression algorithm

Figure 20 shows the results obtained by applying the logistic regression algorithm using dual training. As shown in the previous results, there is no difference in the performances when the number of users is small. However, power control is performed as the number of users increases, and the performance is improved by 1.02 times to 1.36 times, as shown in the graph.

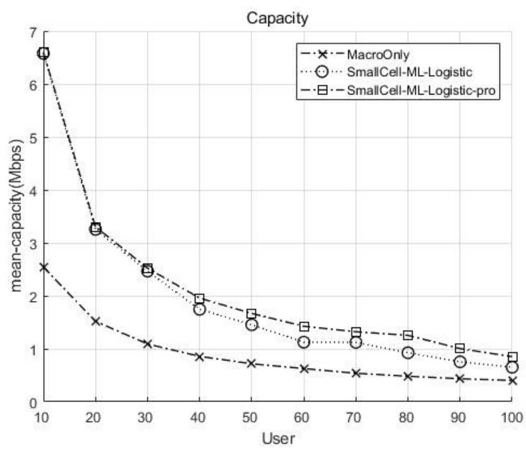

Figure 20: 5G Small Cell Machine Learning Algorithm Average System Total (Logistic Regression Dual Training)

4.3.6 Dual training simulation results for decision tree algorithm

Figure 21 shows the results obtained by applying the decision tree algorithm using the dual training method. As shown in the previous results, there is no difference when the number of users is small. However, effective interference control is performed as the number of users increases, and the overall performance is improved. Compared to the decision tree algorithm, the performance is improved by 0.97 times to 1.23 times.

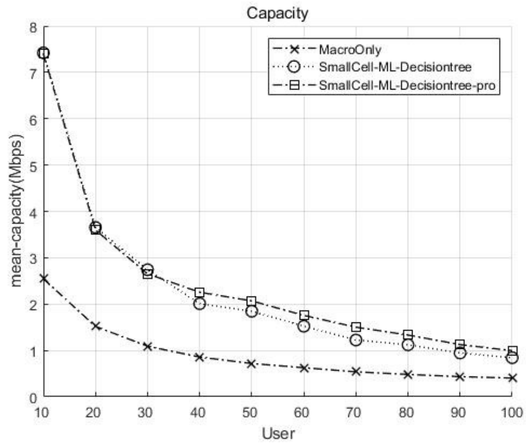

Figure 21: 5G Small Cell Machine Learning Algorithm Average System Total (Decision Tree Dual Training)

5 Conclusions

With the emergence of 4G-LTE systems, it is possible for anyone to easily access video streaming services. As a result, there has been an increase in mobile traffic market. Along with this advancement, there was a significant increase in mobile traffic, and existing 4G systems could no longer handle the increased mobile traffic. The 5G mobile communication system was developed to address this change. The 5G system was developed to accommodate many devices, implement ultra-high-speed data communications, and achieve ultra-low latency. With this technology, frequencies that are higher than those in existing 4G systems are used. However, the use of high frequencies inevitably results in a large path loss, and there are many instances in which user communication cannot be guaranteed. To overcome these problems, research was conducted on the application of small cells to 5G mobile communication systems. The deployment of small cells was implemented to cope with various environments by configuring a large number of small cells in a manner that is similar to a UDN environment. Such an increase in the number of small cells may improve the overall system performance; however, there will inevitably by interference owing to the small base stations located within the base station. Moreover, one base station could become saturated with connected mobile devices. In order to mitigate this, machine learning algorithms were applied in the small cell connection determination phase, and the performance evaluation results of machine learning, including the accuracy, recall, precision, and F1-score, as well as the training time, were analyzed for each algorithm. Although the SVM algorithm showed superiority in three performance categories besides recall, it took longer to perform the training. The training time for the logistic algorithm was 23.7 s, the decision tree algorithm was 6.5 s, and the SVM algorithm was 45.6 s. Hence, there will be penalties when applying the training model in real time in a multilateral environment. Overall, the decision tree algorithm exhibited a similar performance to the SVM algorithm in the areas excluding precision, and it has a very short training time. Therefore, the decision tree algorithm has a superior overall algorithm performance. When each algorithm was applied to the small cell network, the performance improved compared to that of the existing small cell technology. The performance improved by 1.33 to 2.84 times when the SVM algorithm was applied, by 1.20 to 2.17 times when the decision tree algorithm was applied, and by 1.06 to 1.68 times when the logistic regression algorithm was applied. When applying the dual training method that considers power control to improve the performance of the small cell machine learning algorithm system, the performance improved by 1.01 to 1.31 times for the SVM algorithm, 0.97 to 1.23 times for the decision tree algorithm, and 1.02 to 1.36 times for the logistic regression algorithm.

With the commercialization of next-generation mobile communication systems and the future expansion of millimeter-wave services, a larger number of small cells will be deployed as a solution for the use of high frequencies. Moreover, there is the need for further studies in this field when developing mobile communication systems.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References

[1] Jahng, J.H.; park, S.K. (2020). 5G Mobile Traffic Forecast, *Electronics and Telecommunications Trends*, 35(6), 129–136, 2020, DOI:10.22648/ETRI.2020.J.350613

[2] Bang, R.A.; Hong, S.E.; Song, J.T.; Kim, I.K.; Park, A.S.; Lee, M.S.; Jang, S.C. (2013). 5G Mobile Communication Technology trends, *Information & communications magazine*, 30(12), 25–36, 2013.

[3] Ericsson. (2019). Estimated growth of Mobile data Traffic and its distribution, *Ericsson report*, 2019.

[4] Kim, Y.H.; Bae, S.H.; Kim, T.Y. (2021). A Performance Improvement Technique Using Machine Learning in a Simulation-Based 5G-Small Cell Mobile Communication System, *International Journal of Future Generation Communication and Networking*, 14(3), 2021.

[5] Samsung. (2018). Who&How:Making 5G NR Standards, *Samsung White Paper*, 2018.

[6] Sulyman, A.I.; Nassar, A.T.; Samimi, M.K.; MacCartney, G.R.; Rappaport, T.S.; Alsanie, A. (2021). Radio propagation path loss models for 5G cellular networks in the 28 GHz and 38 GHz millimeter-wave bands, *IEEE communications magazine*, 52(9), 78–86, 2014, DOI: 10.1109/MCOM.2014.6894456

[7] Chung, W.G. (2018). Study on Effective 5G Network Deployment Method for 5G Mobile Communication Services, In *The Journal of Korean Institute of Electromagnetic Engineering and Science*, 29(5), 353–358, 2018. DOI: https://doi.org/10.5515/KJKIEES.2018.29.5.353

[8] Patwary, M.; Sharma, S.K.; Chatzinotas, S.; Chen, Y.; Abdel-Maguid, M.; Abd-Alhameed, R.; Noras, J.; Ottersten, B. (2016). Universal Intelligent Small Cell (UnISCell) for next generation cellular networks, *Digital Communications and Networks*, 2(4), 167–174, 2016, https://doi.org/10.1016/j.dcan.2016.09.003

[9] [Online]. Available: https:// telecoms.com/tag/informa-telecoms-media/, Informa Telecoms&Media, SmallCell Market Status, 2013.

[10] 5G.Forum (2018). Small Cell Ecosystem White Paper, 39–40, 2018.

[11] [Online]. Available: http://blog.3g4g.co.uk/2013/04/interference-management-in-hetnets.html, Accesed on 8 Aptil 2013.

[12] Kim, K.; Myung, J.; Seo, J. (2018). Research Trends on Wireless Transmission and Access Technologies Using Deep Learning, *Electronics and Telecommunications Trends*, 33(5), 13–23, 2018, DOI:10.22648/ETRI.2018.J.330502

[13] Woo, Y.C.; Lee, S.Y.; Choi, W.; Ahn, C.W.; Baek, O.K. (2019). Trend of Utilization of Machine Learning Technology for Digital Healthcare Data Analysis, *Electronics and Telecommunications Trends*, 34(1), 98–10, 2019, DOI:10.22648/ETRI.2019.J.340109

[14] Zhang, A.; Lipton, Z.C.; Li, M.; Smola, A.J. (2021). Dive into deep learning, arXiv preprint arXiv:2106.11342.

[15] [Online]. Available: http:// developer.nvidia.com/deep-learning

Cite this paper as:

Kim, Y. H.; Lee, D. Y.; Bae, S. H.; Kim, T. Y. (2021). Improving the Performance of Heterogeneous Network Systems in Machine Learning-based 5G Mobile Communication System, *International Journal of Computers Communications & Control*, 16(6), 4583, 2021.
https://doi.org/10.15837/ijccc.2021.6.4583