

Article

Repetition with Learning Approaches in Massive Machine Type Communications

Li-Sheng Chen ^{1,*}, Chih-Hsiang Ho ², Cheng-Chang Chen ³, Yu-Shan Liang ⁴ and Sy-Yen Kuo ⁵¹ Department of Communications Engineering, Feng Chia University, Taichung 407, Taiwan² Institute for Information Industry, Taipei 106, Taiwan³ Bureau of Standards, Metrology and Inspection, M.O.E.A., Taipei 100, Taiwan⁴ Department of Electronic Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 824, Taiwan⁵ Department of Electrical Engineering, National Taiwan University, Taipei 100, Taiwan

* Correspondence: lschen@fcu.edu.tw

Abstract: In the 5G massive machine type communication (mMTC) scenario, user equipment with poor signal quality requires numerous repetitions to compensate for the additional signal attenuation. However, an excessive number of repetitions consumes additional wireless resources, decreasing the transmission rate, and increasing the energy consumption. An insufficient number of repetitions prevents the successful deciphering of the data by the receivers, leading to a high bit error rate. The present study developed adaptive repetition approaches with the k-nearest neighbor (KNN) and support vector machine (SVM) to substantially increase network transmission efficacy for the enhanced machine type communication (eMTC) system in the 5G mMTC scenario. The simulation results showed that the proposed repetition with the learning approach effectively improved the probability of successful transmission, the resource utilization, the average number of repetitions, and the average energy consumption. It is therefore more suitable for the eMTC system in the mMTC scenario than the common lookup table.

Citation: Chen, L.-S.; Ho, C.-H.; Chen, C.-C.; Liang, Y.-S.; Kuo, S.-Y. Repetition with Learning Approaches in Massive Machine Type Communications. *Electronics* **2022**, *11*, 3649. <https://doi.org/10.3390/electronics11223649>

Academic Editor: Christos J. Bouras

Received: 1 September 2022

Accepted: 19 October 2022

Published: 8 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: massive machine type communications (mMTC); enhanced machine type communication (eMTC); repetition; machine learning; k-nearest neighbor (KNN); support vector machine (SVM)

1. Introduction

Machine type communication (MTC) describes data exchange and communications between machines and is a crucial component in areas such as smart cities, traffic optimization, smart poles, e-medicine, and the industrial Internet of Things (IoT) [1,2]. Following the increase in MTC applications, the number of connected wireless devices has risen dramatically; the number of wireless MTC devices is expected to reach several billion [3]. However, this has created a critical challenge for the devices' network-access capability. Massive MTC (mMTC) is the primary third-generation partnership project (3GPP) fifth-generation (5G) application scenario and technology implemented in IoT devices [4]. The respective technology involves connecting a massive number of components. For each square kilometer of area in a developed region, the machine-machine communications between 1 million devices are expected. Such communications are primarily for small quantities of data with relatively high latency tolerance. The components must be cost-efficient and equipped with long battery lives.

3GPP enhanced MTC (eMTC) and 3GPP narrowband IoT (NB-IoT) are 5G IoT technology standards that address the requirements of 3GPP 5G mMTC application scenarios. Studies have explored the physical characteristics and design goals of eMTC and NB-IoT [5–7]. In the long-term evolution (LTE), the 5G, advanced eMTC, and NB-IoT user equipment (UE) switches from idle mode to connected mode through random access (RA) [8,9];

this process is achieved through a four-way handshake, which involves a preamble transmission, an RA response, a connection request, and a connection resolution. When the UE transmits data in a highly synchronized manner, the signaling capacity of the evolved node B can be easily exceeded, causing severe system congestion [10]. Ali and Hamouda [11] proposed an algorithm involving beehive search and initial synchronization. The simulation results confirmed that the algorithm provides satisfactory network efficacy even for devices with an extremely low signal-to-noise ratio. In [12], a maximum-likelihood detector was applied for initial timed collection in IoT devices. The average detection latency of the detector was half as high as that of the related method, and the energy required for each timed collection decreased by 34%. An approach integrating agent-based modeling and simulation was employed for IoT system efficacy analysis in [13,14]. The agent-based cooperative smart object [15] approach uses OMNeT++ for simulation. Liu et al. [16,17] proposed two novel resource coordination methods that involve bridging dynamic sensing tasks with heterogeneous IoT sensors and controlling the operating cycles of devices to save energy. In [18], eMTC infrastructure coverage efficacy was analyzed and compared with that of LTE technology to reinforce the eMTC coverage at 15 dB.

The design of mMTC technology must fulfill the following four requirements:

a. Coverage: The coverage of mMTC must attain a maximum coupling loss (MCL) of 164 dB [19]. Even when the signal from the transmitter to the receiver is attenuated by as much as 164 dB, the receiver must successfully decipher the packet. In addition, because increasing coverage through repeat transmission reduces the data transmission rate substantially, 5G mMTC coverage must also be successfully implemented at a transmission rate as high as 160 bit/s. Therefore, selecting an appropriate coverage enhancement (CE) level and number of repetitions for the signal quality is critical.

b. UE battery life: Devices with long battery lives are required in 5G MTC applications such as smart electric and water meters. These devices may be installed in environments that impede battery replacement or in which the cost of a battery replacement is excessive. Therefore, mMTC devices must have battery lives of no less than 10 years [20], similar to those of eMTC or NB-IoT devices.

c. Connection density: As the demand for IoT applications increases, the density of 5G IoT devices is expected to reach approximately 1 million devices per km² in developed areas. Therefore, 5G mMTC must support a connection density value as high as 106 devices per km² while maintaining a specific quality of service.

d. Latency: Although most MTC devices have considerable data transmission latency tolerance, the 5G mMTC specifications include a latency tolerance requirement to ensure satisfactory quality of service. Specifically, for each 20-byte application layer packet transmitted by a device, the latency should not exceed 10 s in a channel with an MCL of 164 dB.

Both NB-IoT and eMTC are low-power wide-area network technologies in authorized spectra. NB-IoT is versatile in spectra and can support three modes of deployment. However, eMTC is faster and exhibits a broader range of applications; in a half-duplex system, the uplink and downlink speed of eMTC is 375 kbps, making eMTC applicable for IoT applications for which a medium data rate is required. Regarding peak speed, NB-IoT exhibits almost no mobility because it does not support handover between base stations, whereas eMTC exhibits more favorable mobility.

Several papers have proposed the study and overview of machine learning technologies, applications, and challenges for IoT and 5G networks [21,22]. Ref. [23] proposes the method with a game and transport theoretic approach for a fog load balancing problem. The work provides a feasible and efficient load balancing solution to ensure an optimal job assignment in the fog computing network with the NB-IoT. One study [24] used machine learning techniques that can be applied for the automation of network functions in 5G network slicing. The intelligent station recognition scheme with support vector machines (SVMs) has been proposed to achieve the fine management of stations [25]. A scheduler framework using reinforcement learning has been proposed [26]; the appropriate scheduling strategy is able to maximize user satisfaction, measured in terms of the

distinct quality of the service requirements. Ref. [27] proposes the learning approach to implicitly extract channel features and recover tag symbols using the deep transfer learning (DTL) approach. In [28], a spectrum management architecture was studied and a machine learning-based spectrum decision framework for the IoT network was proposed.

The eMTC requirements support two CE levels (CE Mode A and CE Mode B) [29] and several numbers of repetition, and each piece of UE can select the appropriate CE levels and number of repetitions according to the signal quality. UE with inferior signal quality requires numerous repetitions to compensate for additional signal attenuation, as depicted in Figure 1, which shows the CE level and repetitions for eMTC.

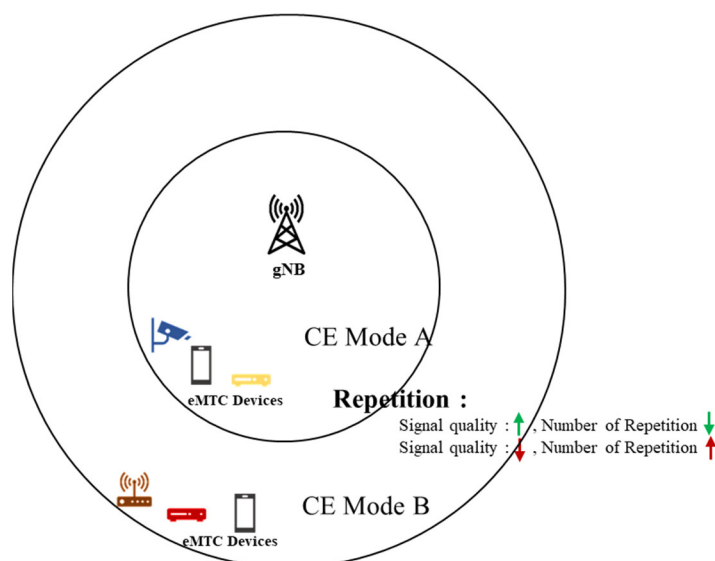


Figure 1. CE level and repetitions for eMTC.

In the eMTC system and mMTC application scenario, due to the main consideration of coverage and power consumption, UE pieces with poor signal quality will use more repetitions to compensate for additional signal attenuation. For the selection of the UE's CE level and repetition times, if they are too high they will waste valuable wireless resources, reduce the transmission rate, and consume more power; if they are too low, the data may not be successfully resolved at the receiving end, resulting in a higher BER (bit error rate). Therefore, we propose the learning adaptive repetition approaches based on k-nearest neighbor (KNN) and the support vector machine (SVM) for an eMTC system in an mMTC application scenario. KNN is easy to implement, has high accuracy, and is insensitive to outliers. When KNN selects appropriate training parameters, the KNN approach can achieve high discrimination accuracy. SVM can avoid the neural network structure selection and local minima problem. SVM does not have a general solution to nonlinear problems; therefore, we significantly choose the kernel function to handle it. For the selection of the CE level and repetition times, we adopted KNN and SVM to propose a learning-based repetition approach, which effectively saves energy efficiency and improves overall network transmission performance. In the present study, a repetition with the learning approach (RLA) was proposed and developed with the aim of substantially increasing the network transmission efficacy. The main contributions of this paper are summarized as follows: (1) we study the problem of repetition number selection for the eMTC system in the mMTC scenario, where the main goal is to minimize the average energy consumption and block error rate (BLER) and maximize the successful transmission probabilities of UE pieces and resource utilization. (2) Due to the limited UE energy efficiency and slave resources of eMTC, we provided the easy-to-implement and learning-based policy for repetition number selection to ensure that eMTC saves energy efficiency and improves transmission efficiency. (3) We adopted the KNN approach and selected

the appropriate training parameters meaningfully so that the KNN approach could be easily implemented and could have high accuracy. (4) We adopted the SVM approach to avoid the neural network structure selection and local minima problem and chose the kernel function significantly. (5) Finally, we conducted an extensive numerical analysis to evaluate the performance of the proposed repetition number selection approaches, and the simulation results show that our proposed approaches (RLA-KNN and RLA-SVM) are more efficient than the common lookup table (LUT) method. The proposed repetition with the learning approach can effectively improve the probability of successful transmission, resource utilization, average number of repetitions, and average energy consumption. The remainder of this paper is organized as follows. Section 2 describes the eMTC system. The proposed approaches are presented in Section 3. In Section 4, we evaluate the performance of our proposed approaches through simulation. Finally, Section 5 provides the conclusions.

2. eMTC System

eMTC technology involves enhancing and customizing mMTC according to the application requirements, such as repetition, scheduling, discontinuous reception, control channels, and system information block (SIB). The eMTC UE monitors the same master information blocks (MIBs), the scheduling period of which is 40 ms. With eMTC, however, within 40 ms, the MIB information is transmitted at the 0th, 18th, and 19th time slots in each radio frame. As such, the coverage of a physical broadcast channel is enhanced through repetition.

In the 3GPP protocol, eMTC independently defines the SIB information; that is, SIB-BR is composed of SIB1-BR and other numbers of SIB information. SIB-BR features a time-frequency position different from that of LTE. SIB1-BR is transmitted on a physical downlink shared channel (PDSCH) at a fixed period of eight radio frames (80 ms). Within each period, SIB1-BR is repeatedly transmitted 4, 8, or 16 times according to the designation by an MIB.

SIB1-BR is transmitted without a controlled schedule through frequency hopping between NBs. The pattern of the frequency hopping is related to the system bandwidths and physical-layer cell identities. The other SIB-BR information is transmitted in the PDSCH; the system information is transmitted without controlled schedules. The schedule information of the system information, such as the time-frequency locations, modulation coding scheme levels, and number of repetitions, is determined in SIB1-BR.

When changes occur in the broadcast information, except for SIB10-BR, SIB11-BR, SIB12-BR, and SIB14-BR, an update is announced through paging or DCI6-2 during the period of change, and new broadcast information is transmitted in the subsequent period of change. The same procedure applies when changes occur in SIB10-BR, SIB11-BR, SIB12-BR, and SIB14-BR.

Unlike NB-IoT, eMTC supports the switching of connection statuses and uninterrupted service transmission when the UE moves across regions. Therefore, eMTC is applicable for services such as voice calls and logistics tracking. The energy-saving function of eMTC is achieved through extension discontinuous reception, power saving mode, and the reduction in the periodic position update frequency. The 3GPP system extends the time of the periodic position update procedure to reduce the periodic position update frequency of the UE, thereby mitigating the signal load in the networks and the power expense in the UE. With the power saving mode of the 3GPP R12 system, the UE timer enters a deep sleep mode after the UE's tracking area is updated and its attach procedure is completed. In the deep sleep mode, the UE does not monitor paging, and the radio transceiver unit is switched off to save significantly more power than that saved in the idle mode. The UE remains registered in the network in the deep sleep mode without the requirement of rerunning the attach procedure or reestablishing the packet data network connection.

The physical layer of eMTC is redesigned through 3GPP, and the eMTC signal coverage is enhanced through the repetition mechanism in the physical channel. In downlink, eMTC employs neither a physical control format indicator channel nor a physical hybrid automatic repeat request indicator channel. The number of signals transmitted in the 40 ms period is increased in the physical broadcast channel. Rather than the conventional

LTE physical downlink control channel, an MTC physical downlink control channel is employed for a maximal repetition number of 256. A maximum of six RBs can be applied by a UE unit in a PDSCH. The modulation coding scheme level and the maximal number of repetitions are restricted; specifically, the maximal number of repetitions is set as 2048. Similarly, a maximum of six RBs can be applied by a UE unit in a physical uplink shared channel (PUSCH). The modulation coding scheme level and the maximal number of repetitions are restricted. In particular, the maximal number of repetitions is set as 2048. In a physical random access channel, a maximum of six RBs can be applied by a UE unit, and a maximum of 128 repetitions can be performed. In a physical uplink control channel (PUCCH), a maximum of 32 repetitions can be performed. The coverage modes of the UE connected to a network can be categorized into two types, namely CE mode A, which indicates satisfactory coverage with no or few repetitions, and CE mode B, which involves more repetitions. In CE mode A, the maximal number of repetitions is 32 in a PDSCH or a PUSCH and 8 in a PUCCH; in CE mode B, the maximal number of repetitions is 2048 in a PDSCH or a PUSCH and 32 in a PUCCH.

3. Repetition with Learning Approaches

Most repetition approaches involve determining the number of repetitions according to the one-dimensional channel quality measured in a data table generated through a simulation. This method is strongly reliant on specific channel models, and errors in the channel quality measurement can lead to the selection of an incorrect number of repetitions. Moreover, a large-dimension data table requires considerable memory space. With the future trend of repetition technology development, one-dimensional channel quality indicators can no longer comprehensively represent the conditions of the channels in complex systems. This is because the link efficacy is related to numerous system parameters, and this can lead to erroneous channel quality assessment and inhibited system efficiency. The accuracy of channel quality indicators can be improved through an increase in the dimension of the channel quality. Machine learning provides a system with high-dimensional channel quality indicators that take into account changes in the environment. The relationship between channel quality and link reliability can be identified through examining the historical records in data transmission, thereby attaining an accurate prediction of the number of repetitions required according to the channel quality measurement results. Conventional repetition approaches are based on inquired data tables and cannot be updated according to system environments; this reduces their flexibility. Machine learning enables a system to learn and adapt to changes in an environment and flexibly adjust its channel quality selection standards, thereby accurately predicting the number of repetitions required. Accordingly, research on the implementation of machine learning to predict the number of uplink repetitions in eMTC is paramount for system efficacy enhancement.

K-nearest neighbor (KNN) is a supervised learning approach in which a particular sample is considered to belong to a specific class if most of the *k*-nearest neighbors to it in its eigenspace also belong to said class [30–32]. In particular, a non-classified sample is categorized to the class in which its nearest classified neighbors belong. When new samples enter a group of classified samples, their distances from the training data are calculated, and the *k*-nearest samples are determined. Subsequently, the class with the greatest number of neighbors among the *k*-selected data is determined, and the new sample is classified accordingly.

In the training phase, according to the feature set acquired from the signal-to-interference-plus-noise ratio (SINR) of each subcarrier, number *i*, which corresponds to repetition number, is selected to optimize the data rate and categorize the training set. All of the elements in each feature set must satisfy the optimized class, which maximizes the number of data transmitted, subject to the limitations of the block error rate.

$$\arg \max_i \{RN_i: BLER_i < H\} \quad (1)$$

RLA-KNN approach training requires a training set with SINR feature vectors. Each vector is assigned to a number i according to equation (1). In RLA-KNN, when the training sample set and the K value and distance are determined, the class of any new sample can be determined. $U = \{u_1, u_2, \dots, u_n\}$, representing the eigenvector training set, and $S = \{s_1, s_2, \dots, s_n\}$, representing the class to which each eigenvector corresponds. RLA-KNN approach training is performed to identify all the vectors in a feature set and the approximated repetition approach corresponding to them.

Let the training sample set TSS in KNN be

$$\text{TSS} = \{(s_1, u_1), (s_2, u_2), \dots, (s_n, u_n)\} \quad (2)$$

The sample to be classified, x , is imported. The distance between x and all other samples in the training set can be calculated as follows: $L(x, s_j), j = 1, 2, \dots, n$. According to the distance data, K samples in training set that are closest to x are determined and referred to as $\text{Near}_k(x)$.

Figure 2 illustrates the RLA using KNN (RLA-KNN), where k is the number of comparisons, and $L(x, s_j)$ is a measured distance.

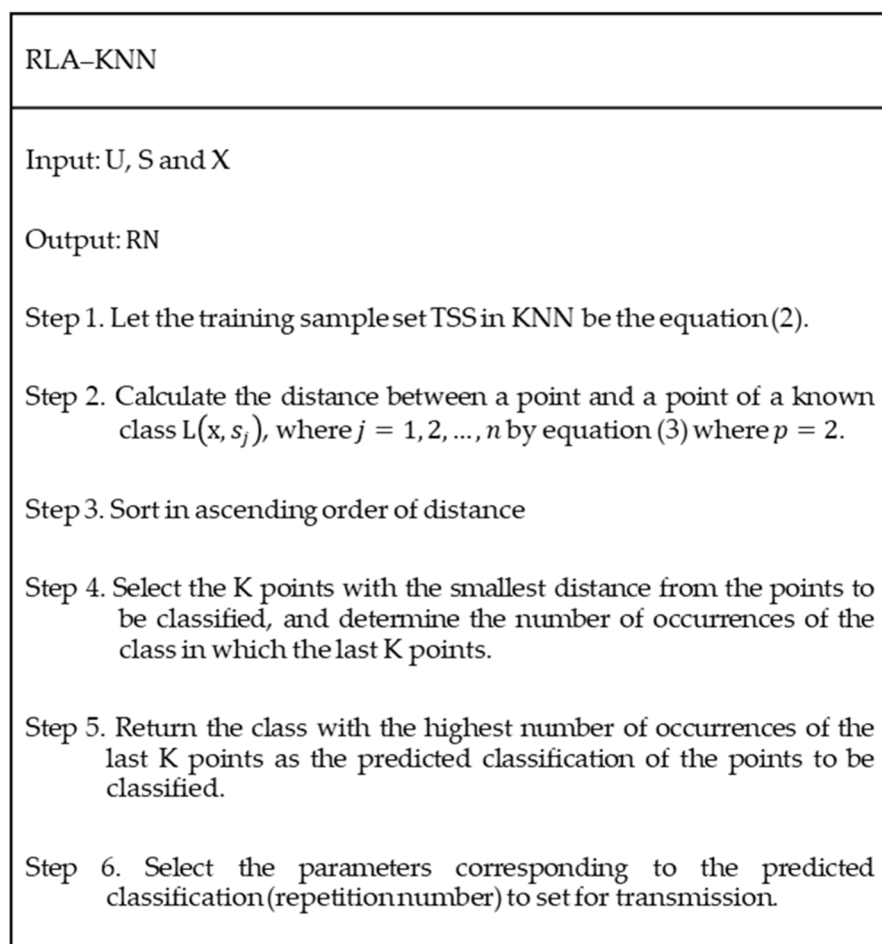


Figure 2. RLA-KNN approach.

In the KNN approach, distance measurement is the most effective approach to determining the similarity between two samples. Commonly used distance measures include the Euclidean distance, Manhattan distance, Chebyshev distance, and Minkowski distance.

Let the eigenspace X be an n -dimension real vector space (R^n): $x_i, x_j \in X$; $x_i = (x_i^1, x_i^2, \dots, x_i^n)^T$; and $x_j = (x_j^1, x_j^2, \dots, x_j^n)^T$. The distance between x_i and x_j is defined as [33]:

$$L_p(x_i, x_j) = \left(\sum_{l=1}^n |x_i^{(l)} - x_j^{(l)}|^p \right)^{1/p} \quad (3)$$

When $p \geq 1$, the Minkowski distance is derived. When $p = 1$, the Manhattan distance is derived. When $p = 2$, the Euclidean distance is obtained. When $p = \infty$, the Chebyshev distance is derived. In the RLA-KNN, we set the p to 2, using distance measures by Euclidean distance.

The time complexity of the prediction for the RLA-KNN approach is $O(N)$, and the time complexity of sorting for the RLA-KNN approach is $O(N \log N)$. KNN approaches are generally inapplicable for large-scale communication systems because of the complex calculation procedure involved. Therefore, SVMs are implemented to identify the repetition approach according to high-dimensional channel quality measurement results. SVM is a commonly employed machine learning method, which is generally modeled as a convex quadratic programming problem [34,35]. Therefore, an SVM is simpler to execute than the KNN approach.

Figure 3 depicts the RLA using an SVM (RLA-SVM) approach, in which a repetition approach is selected according to the BLER in the data.

RLA-SVM
Input: X, Y
Output: RN
Step 1. Let the training sample set in SVM be the equation (4).
Step 2. The new dataset generated by the equation (5) is linearly separable .
Step 3. Choose a function: $k(x_1, x_2) = \langle x_1, x_2 \rangle^2$, use it to replace the inner product calculation of SVM.
Step 4. For the new dataset after dimension expansion, use equation (6) and equation (7) to predict prediction y.
Step 5. Select the parameters corresponding to the predicted classification (repetition number) to set for transmission.

Figure 3. RLA-SVM approach.

For a set of data X and label Y , the task of SVM is to find a set of parameters such that $x^{\theta^T} = threshold$, and the samples with $x^{\theta^T} < threshold$ are judged as negative samples, and the samples with $x^{\theta^T} > threshold$ are judged as positive samples. The sample X is a two-dimensional data set of channel quality:

$$X = \begin{bmatrix} x_1^{(1)} & x_2^{(1)} \\ x_1^{(2)} & x_2^{(2)} \\ \vdots & \vdots \\ x_1^{(m)} & x_2^{(m)} \end{bmatrix} \quad (4)$$

This dataset X is linearly inseparable on the two-dimensional plane, and the new dataset generated by a transformation function $\phi(x)$ is linearly separable:

$$\phi(x_1, x_2) = (x_1^2, \sqrt{2}x_1, x_2, x_2^2) \quad (5)$$

For the new dataset after the dimension increase, the prediction y for repetition number can be expressed as:

$$\begin{aligned} y_{\text{prediction}} &= \sum_{i=1}^m \lambda_i y^{(i)} \langle \phi(\hat{x}), \phi(x^{(i)}) \rangle \\ &= \sum_{i=1}^m \lambda_i y^{(i)} \langle ((\hat{x}_1^2, \sqrt{2}\hat{x}_1, \hat{x}_2, \hat{x}_2^2)), (x_1^{(i)2}, \sqrt{2}x_1^{(i)}, x_2^{(i)}, x_2^{(i)2}) \rangle \\ &= \sum_{i=1}^m \lambda_i y^{(i)} (\hat{x}_1^2 x_1^{(i)2} + 2\hat{x}_1 x_1^{(i)} \hat{x}_2 x_2^{(i)} + \hat{x}_2^2 x_2^{(i)2}) \\ &= \sum_{i=1}^m \lambda_i y^{(i)} \langle (\hat{x}_1, \hat{x}_2), (x_1^{(1)}, x_2^{(2)}) \rangle^2 \\ &= \sum_{i=1}^m \lambda_i y^{(i)} \langle \hat{x}_1, x_1^{(i)} \rangle^2 \end{aligned} \quad (6)$$

Through the above transformation, after the data set is upscaling, the calculation of SVM training and prediction can actually be converted into the calculation of the original feature. We choose a function: $k(x_1, x_2) = \langle x_1, x_2 \rangle^2$, use it to replace the inner product calculation of SVM:

$$\hat{y} = \begin{cases} +1, & \sum_{i=1}^m \lambda_i y^i k(\hat{x}, x^i) \geq +1 \\ -1, & \sum_{i=1}^m \lambda_i y^i k(\hat{x}_1, x^i) \leq +1 \end{cases} \quad (7)$$

The training and prediction process of this SVM model is equivalent to doing it in a high-dimensional space, which achieves the purpose of linearly dividing the data set, and does not add complex operations, among which is the kernel function.

The loss function of SVM is:

$$\begin{aligned} Loss &= \frac{1}{2} \|\theta\|_2 + \sum_i \max(0, 1 - y\hat{y}) \\ &= \frac{1}{2} \|\theta\|_2 + \sum_i \max(0, 1 - y(x\theta + \theta_0)) \end{aligned} \quad (8)$$

The predicted result corresponds to a repetition number. Select the parameters corresponding to the predicted classification (repetition number) to set for transmission.

4. Performance Evaluation

The RLA was applied to evaluate efficacy. The performances for RLA-KNN and RLA-SVM were analyzed subject to various parameters for a single piece of UE, evaluated in a simulated environment and compared with those of the common lookup tables (LUTs) from the eMTC system optimization. This analysis included the successful transmission probabilities of the UE, the average numbers of the repetitions, the resource utilization, and the average energy consumption.

The traffic model defined in 3GPP TR 36.763 [36] was used as the traffic model in the simulation, where the number of eMTC devices was set to 20,000–200,000 per sector; the

average number of uplink traffic reports was set to 20/s; and the payload size of each device ranged from 20 to 200 bytes. In the simulation, each point represented the average of 10,000 samples, each of which was acquired between the first uplink transmission and the end of the last uplink transmission (i.e., the observation interval). The relevant simulation environment parameters are illustrated in Table 1. The indicators used to evaluate the performance of the proposed approach were as follows: (1) BLER versus signal to interference plus noise ratio (SINR) for various k values in RLA-KNN using a single piece of UE; (2) throughput versus SINR for the various k values in RLA-KNN using a single piece of UE; (3) BLER versus SINR with the RLA-SVM-radial basis function (RBF), RLA-SVM-linear, and RLA-KNN for a single piece of UE; (4) throughput versus SINR with RLA-SVM-RBF, RLA-SVM-linear, and RLA-KNN for a single piece of UE; (5) the successful transmission probabilities of the UE pieces; (6) the average number of repetitions; (7) the resource utilization; and (8) the average energy consumption.

Table 1. Simulation Parameters.

Parameter	Assumption
Sector per cell	3
Fractional frequency reuse	Hard
Allocation size	12 tones @ 15 KHz
Resource Unit (RU)	1 ms
eMTC Bandwidth	6 PRBs (1.08 MHz)
Channel Model	Typical Urban (TU)
Payload size	20~200
Number of eMTC UE pieces per sector	20,000~200,000
DL Antenna Configuration	gNB: 2Tx/4Tx, UE: 1Rx
Uplink Antenna Configuration	gNB: 2Rx/4Rx, UE: 1Tx
Doppler Spread	1 Hz
gNB Tx power	43 dBm
UE Tx power	20 dBm

The RLA-KNN repetition approach consists of a training phase and a test phase. The training phase requires a training set, each of which corresponds to a unique SINR sequence and is assigned to a repetition number. During the training phase, all the vectors in the training set must undergo all repetitions for accurate classification of the set to the repetition number. The vectors are then classified according to the BLER value of each repetition.

As shown in Figures 4 and 5, a higher k leads to higher system efficacy. However, after k exceeds a certain value, the system efficacy begins to drop. Ordinarily, a higher k limits the effect of erroneous classification on the training results more effectively, yielding a smaller classification error during the test phase and higher system efficacy. However, because of a limit in the size of the training set, the error rate starts to increase after k exceeds a certain value, lowering system efficacy. As shown in the simulation results, the system efficacy was maximized when k was 30; so, this value was employed in the follow-up simulation.

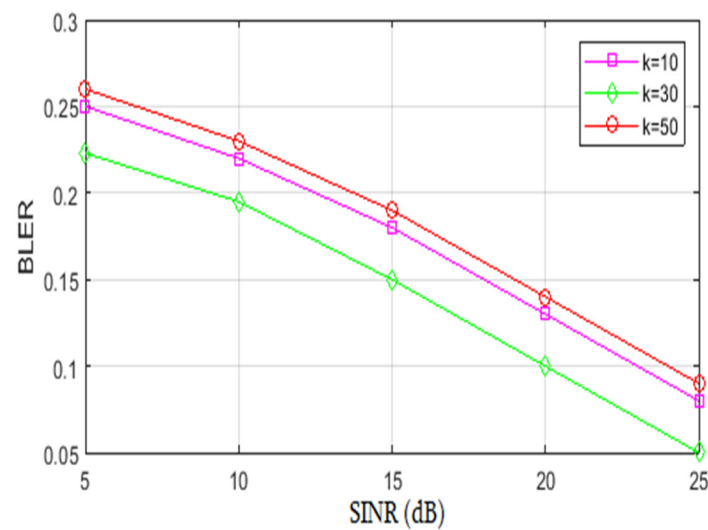


Figure 4. BLER versus SINR for various k values in RLA-KNN.

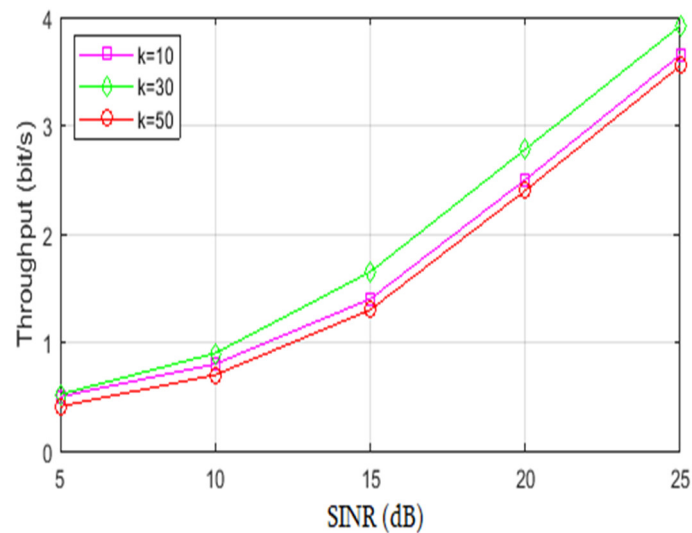


Figure 5. Throughput versus SINR for the various k values in RLA-KNN.

Accordingly, when the number of data in the system is high, the KNN approach requires a complicated calculation procedure and cannot be unsuitably applied in actual network communications without dimensionality reduction. Therefore, the SVM approach is relatively flexible because it can employ numerous kernel functions, reducing the complexity of the calculation procedure considerably. In an RLA-SVM repetition approach, to classify a new sample (the SINR set of a subcarrier), the interference function $h_m(x)$ of each repetition and the BLER of approximately $r_m(h_m(x))$ corresponding to each of the said functions must be identified. We compared the efficacy of the RLA-SVM approach with that of the RLA-KNN approach.

As depicted in Figures 6 and 7, for RLA-SVM the linear kernel function is nearly as efficient as the Gaussian RBF kernel. In particular, the RBF kernel function is slightly more efficient than the linear kernel function, but the RBF kernel function involves a more complicated calculation procedure and requires a larger space for data storage. Moreover, the RLA-SVM repetition approach involves considerably lower calculation and time complexity than does the RLA-KNN repetition approach, even though the two are almost equally efficient in data transmission. Therefore, the RLA-SVM repetition approach is more applicable for an actual eMTC system than the RLA-KNN repetition approach.

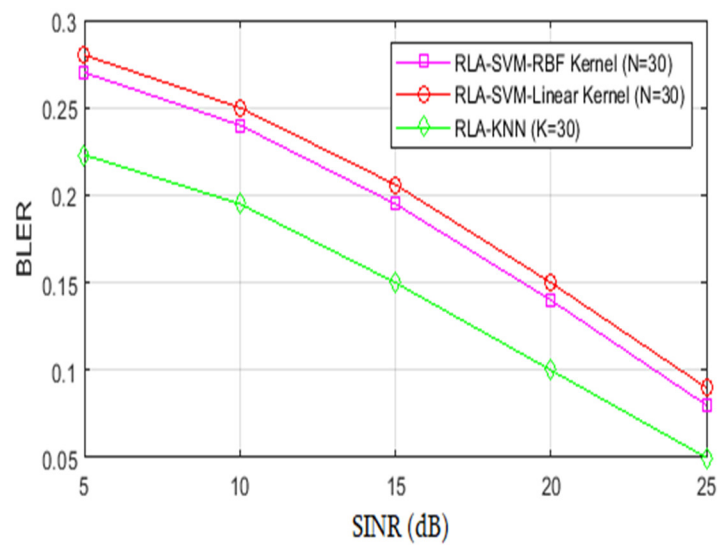


Figure 6. BLER versus SINR for RLA-SVM-RBF, RLA-SVM-linear, and RLA-KNN.

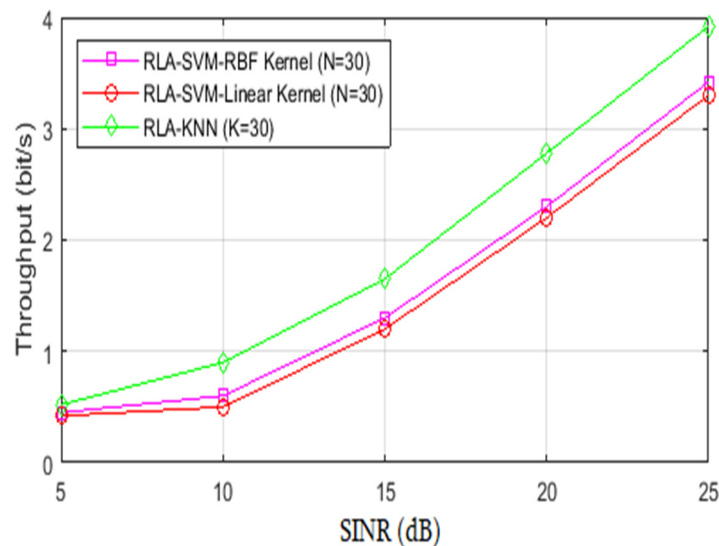


Figure 7. Throughput versus SINR for RLA-SVM-RBF, RLA-SVM-linear, and RLA-KNN.

After analysis of the performance of RLA-KNN and RLA-SVM subject to various parameters for a single piece of UE, to optimize the eMTC system we compare the performances of RLA-KNN, RLA-SVM, and LUT. This analysis includes the probability of successful transmission, the average number of repetitions, resource utilization, and average energy consumption.

The successful transmission probabilities for RLA-KNN, RLA-SVM, and the LUT are shown in Figure 8. The probability of successful transmission gradually decreases as the intensity of the number of pieces of UE per sector increases (i.e., as more users attempt access). Failure probability is observed for high intensity with the use of the LUT method. Apart from collision, the increased failure rate is due to the interference and radio channel effects, caused by the presence of many pieces of UE with wide coverage close to the eNB, which affects such UE pieces farther away in the same coverage zone. The probabilities of successful transmission for RLA-KNN and RLA-SVM are higher than for the LUT method because the repetition number selection policy is considered to be an interference factor from the training data.

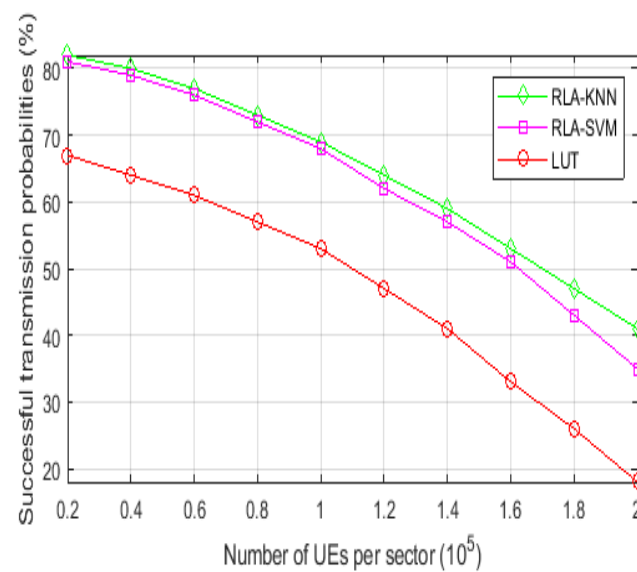


Figure 8. Probability of successful transmission according to UE numbers.

Figure 9 presents the average numbers of repetitions in the RLA-KNN, RLA-SVM, and LUT methods. With the RLA-KNN and RLA-SVM approaches, a significant reduction in repetition is observed. The repetition for the LUT method is noted to be greater than for the RLA-KNN and RLA-SVM methods. In many cases, the LUT method uses excessive numbers of repetitions to achieve successful transmission. In RLA-KNN and RLA-SVM, the number of repetitions required to achieve successful transmission is significantly reduced by the training data and learning processes.

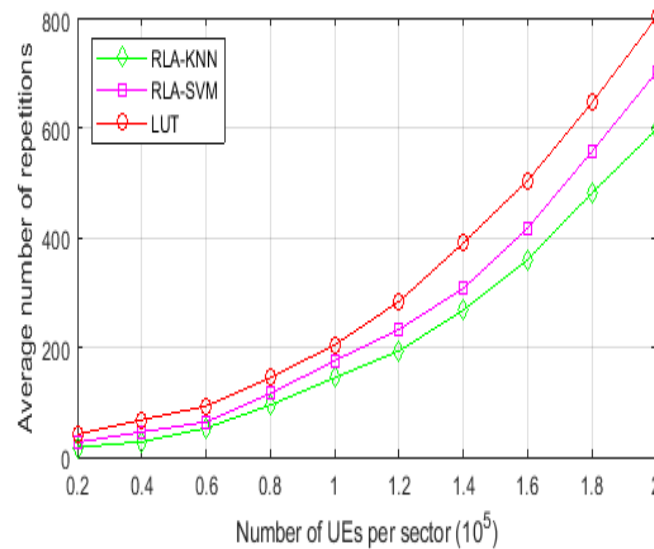


Figure 9. Average number of repetitions.

The resource utilization for RLA-KNN, RLA-SVM, and the LUT is presented in Figure 10. The resource utilization for the three methods is seen to decrease as the number of UE pieces per sector increases. RLA-KNN and RLA-SVM also exhibit higher resource utilization than the LUT because more resources are utilized in the same time period if a shorter repetition is adopted. The result indicates that a smaller number of repetitions is associated with greater resource utilization.

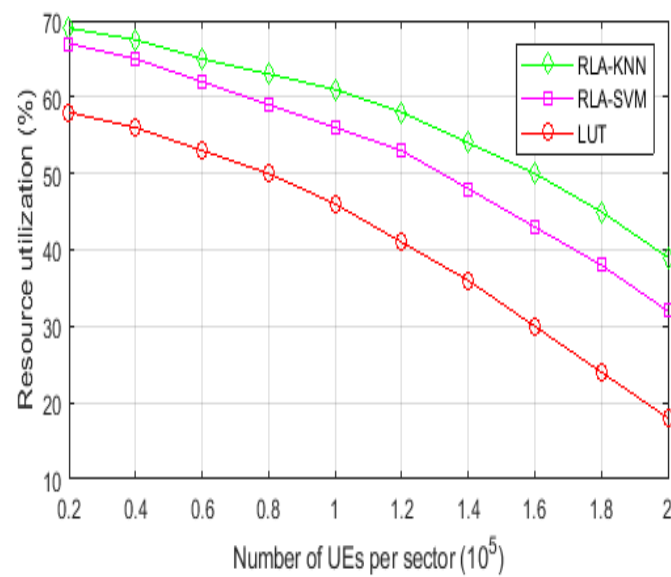


Figure 10. Resource utilization.

Figure 11 presents the average energy consumption per sector in the transmission mode by all users. This decreases the number of repetitions and, thus, power consumption. With the suitable reduction in repetition using the RLA-KNN and RLA-SVM methods, energy consumption is also significantly reduced. A significant reduction in energy consumption is achieved for the uplink compared with for the LUT method. More frequent transmission increases power consumption quickly, which is impractical for IoT devices. More efforts are required to reduce the power consumption of the radio-frequency modules, such as the number of repetitions, under the satisfaction of successful transmission.

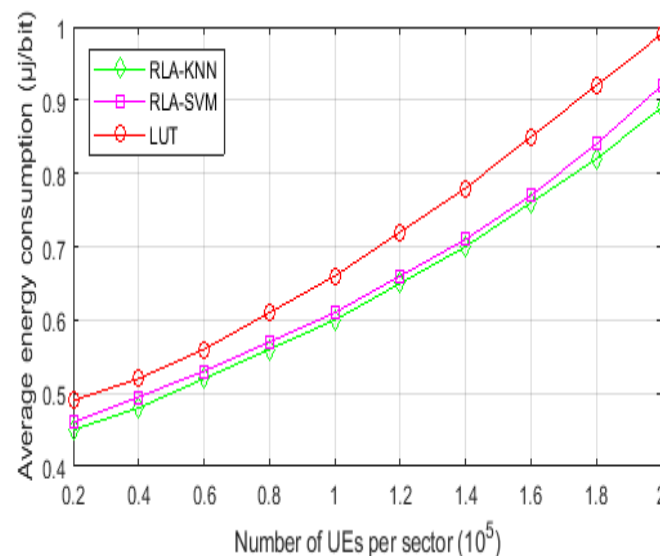


Figure 11. Average energy consumption.

5. Conclusions

An insufficient number of repetitions prevents the successful deciphering of the data by the receivers, leading to a high bit error rate. Excessively high repetitions of UE pieces lead to the wastage of valuable wireless resources in 3GPP mMTC. Therefore, in the present study, adaptive repetition approaches with machine learning were developed to substantially increase the network transmission efficacy for mMTC systems in mMTC. The simulation results show that the proposed RLA could effectively improve the transmission

probabilities, the resource utilization, the average number of repetitions, and the average energy consumption. The proposed RLA is more suitable than the common LUT for the eMTC system in mMTC. In future work, we will adopt online deep learning approach for 6G or open radio access network (O-RAN) AI architecture and specific use cases. We will enable the approach to learn effectively in special communication scenarios with difficult-to-obtain training samples and have more appropriate depth and accuracy of learning.

Author Contributions: Conceptualization, L.-S.C., C.-H.H., C.-C.C., Y.-S.L. and S.-Y.K.; Writing—original draft, L.-S.C. and C.-H.H.; Writing—review & editing, C.-C.C., Y.-S.L. and S.-Y.K. All authors have read and agreed to the published version of the manuscript.

Funding: The original research work presented in this paper is partly related to BSMI under contract number 1D201101221-134 awarded by BSMI Taiwan.

Institutional Review Board Statement: Not applicable.

Acknowledgments: The authors would like to thank the BSMI (Bureau of Standards, Metrology and Inspection, Taiwan) research group and MOST for their technical support.

Conflicts of Interest: The authors declare no conflict of interest

References

- Service Requirements for Machine-Type Communications (Release 17), Document 3GPP TS 22.368 V17.0.0, April 2022. Available online: https://www.etsi.org/deliver/etsi_ts/122300_122399/122368/17.00.00_60/ts_122368v170000p.pdf (accessed on 18 October 2022).
- Verma, R.; Prakash, A.; Agrawal, A.; Naik, K.; Tripathi, R.; Khalifa, T.; Alsabaan, M.; Abdelkader, T.; Abogharaf, A. Machine-to-machine (M2M) communications: A survey. *J. Netw. Comput. Appl.* **2016**, *66*, 83–105.
- Qipeng, S.; Nuaymi, L.; Lagrange, X. Survey of radio resource management issues and proposals for energy-efficient cellular networks that will cover billions of machines. *EURASIP J. Wirel. Commun. Netw.* **2016**, *2016*, 140.
- Bockelmann, C.; Pratas, N.K.; Wunder, G.; Saur, S.; Navarro, M.; Gregoratti, D.; Vivier, G.; de Carvalho, E.; Ji, Y.; Stefanovic, C.; et al. Towards massive connectivity support for scalable mMTC communications in 5G networks. *IEEE Access* **2018**, *6*, 28969–28992.
- Mahjoubi, A.E.; Mazri, T.; Hmina, N. NB-IoT and eMTC: Engineering results towards 5G/IoT mobile technologies. In Proceedings of the 2018 International Symposium on Advanced Electrical and Communication Technologies (ISAECT, Rabat, Morocco, 21–23 November 2018; pp. 1–7.
- 5G.; Study on Scenarios and Requirements for Next Generation Access Technologies (Release 17), Document 3GPP TR 38.913 V17.0.0, April 2022. Available online: <https://standards.iteh.ai/catalog/standards/etsi/17425546-b012-499b-8730-15e038e8186d/etsi-tr-138-913-v17-0-0-2022-05> (accessed on 18 October 2022).
- El Mahjoubi, A.; Mazri, T.; Hmina, N. NB-IoT, eMTC, eMIMO and Massive CA: First Africa Engineering Experimental Results Towards the Delivery of 5G/IoT Smart Technologies Applications. *J. Commun.* **2019**, *14*, 216–222.
- Service Accessibility (Release 17), Document 3GPP TS 22.011 V17.5.0, December 2021. Available online: https://www.etsi.org/deliver/etsi_ts/122000_122099/122011/17.05.00_60/ts_122011v170500p.pdf (accessed on 18 October 2022).
- Physical Channels and Modulation (Release 17), Document 3GPP TS 36.211 V17.0.0, January 2022. Available online: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2425> (accessed on 18 October 2022).
- Radio Resource Control (RRC); Protocol specification (Release 17), Document 3GPP TS 36.331 V17.0.0, April 2022. Available online: <https://www.tech-invite.com/3m36/tinv-3gpp-36-331.html> (accessed on 18 October 2022).
- Ali, A.; Hamouda, W. On the cell search and initial synchronization for NB-IoT LTE systems. *IEEE Commun. Lett.* **2017**, *21*, 1843–1846.
- Kroll, H.; Korb, M.; Weber, B.; Willi, S.; Huang, Q. Maximumlikelihood detection for energy-efficient timing acquisition in NB-IoT. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), San Francisco, CA, USA, 19–22 March 2017; pp. 1–5.
- Fortino, G.; Gravina, R.; Russo, W.; Savaglio, C. Modeling and Simulating Internet-of-Things Systems: A Hybrid Agent-Oriented Approach. *Comput. Sci. Eng.* **2017**, *19*, 68–76.
- Fortino, G. Agents meet the IoT: Toward ecosystems of networked smart objects. *IEEE Syst. Man Cybern. Mag.* **2016**, *2*, 43–47.
- Fortino, G.; Russo, W.; Savaglio, C. Simulation of Agent-oriented Internet of Things Systems. In Proceedings of the 17th Workshop “From Objects to Agents, Catania, Italy, 29–30 July 2016; pp. 8–13.
- Liu, C.H.; Fan, J.; Branch, J.W.; Leung, K.K. Toward QoI and energy-efficiency in Internet-of-Things sensory environments. *IEEE Trans. Emerg. Topics Comput.* **2014**, *2*, 473–487.
- Liu, C.H.; Fan, J.; Hui, P.; Wu, J.; Leung, K.K. Toward QoI and energy efficiency in participatory crowdsourcing. *IEEE Trans. Veh. Technol.* **2015**, *64*, 4684–4700.
- Ratasuk, R.; Bhatoolaul, D.; Mangalvedhe, N. Performance Analysis of Voice over LTE Using Low-Complexity eMTC Devices. In Proceedings of the 2017 IEEE 85th Vehicular Technology Conference (VTC Spring), Sydney, NSW, Australia, 4–7 June 2017.
- Böcker, S.; Arendt, C.; Jörke, P.; Wietfeld, C. LPWAN in the Context of 5G: Capability of LoRa WAN to Contribute to mMTC. In Proceedings of the 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 15–18 April 2019.

20. Soussi, M.E.; Zand, P.; Pasveer, F.; Dolmans, G. Evaluating the performance of eMTC and NB-IoT for smart city applications. In Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 20–24 May 2018; pp. 1–7.
21. Sharma, S.K.; Wang, X. Towards massive machine type communications in ultra-dense cellular IoT networks: Current issues and machine learning-assisted solutions. *IEEE Commun. Surv. Tutor.* **2019**, *22*, 426–471.
22. Calabrese, F.D.; Wang, L.; Ghadimi, E.; Peters, G.; Hanzo, L.; Soldati, P. Learning Radio Resource Management in RANs: Framework Opportunities and Challenges. *IEEE Commun. Mag.* **2018**, *56*, 138–145.
23. Abedin, S.F.; Bairagi, A.K.; Munir, M.S.; Tran, N.H.; Hong, C.S. Fog Load Balancing for Massive Machine Type Communications: A Game and Transport Theoretic Approach. *IEEE Access* **2019**, *7*, 4204–4218.
24. Kafle, V.P.; Fukushima, Y.; Julia, P.M.; Miyazawa, T. Consideration On Automation of 5G Network Slicing with Machine Learning. In Proceedings of the 2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K), Santa Fe, Argentina, 26–28 November 2018.
25. Liu, Z.; Dai, Z.; Yu, P.; Jin, Q.; Du, H.; Chu, Z.; Wu, D. Intelligent station area recognition technology based on NB-IoT and SVM. In Proceedings of the IEEE 28th International Symposium on Industrial Electronics, Vancouver, BC, Canada, 12–14 June 2019.
26. Comsa, I.S.; Domenico, A.D.; Ktenas, D. QoS-driven scheduling in 5G radio access networks—A reinforcement learning approach. In Proceedings of the GLOBECOM 2017–2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–7.
27. Liu, C.; Wei, Z.; Ng, D.W.K.; Yuan, J.; Liang, Y.-C. Deep transfer learning for signal detection in ambient backscatter communications. *IEEE Trans. Wireless Commun.* **2021**, *20*, 1624–1638.
28. Li, L.; Ghasemi, A. IoT Enabled Machine Learning for an Algorithmic Spectrum Decision Process. *IEEE IoT J.* **2019**, *6*, 1911–1919.
29. Reddy, M.P.; Santosh, G.; Kumar, A.; Kuchi, K. Improved Physical Downlink Control Channel for 3GPP Massive Machine Type Communications. In Proceedings of the International Conference on Communication Systems and Networks, Gurgaon, India, 1–2 December 2018; pp. 1–25.
30. Daniels, R.C.; Caramanis, C.; Heath, R.W. A supervised learning approach to adaptation in practical MIMO-OFDM wireless systems. In Proceedings of the IEEE GLOBECOM 2008–2008 IEEE Global Telecommunications Conference, New Orleans, LA, USA, 30 November–4 December 2008; pp. 1–5.
31. Daniels, R.C.; Caramanis, C.; Heath, R.W., Jr. Adaptation in convolutionally coded MIMO-OFDM wireless systems through supervised learning and SNR ordering. *IEEE Trans. Veh. Technol.* **2010**, *59*, 114–126.
32. Puljiz, Z.; Park, M.; Heath, R.W., Jr. A machine learning approach to link adaptation for SC-FDE system. In Proceedings of the 2011 IEEE Global Telecommunications Conference-GLOBECOM, Houston, TX, USA, 5–9 December 2011; pp. 1–5.
33. Sun, X.-Q.; Chen, Y.-J.; Shao, Y.-H.; Li, C.-N.; Wang, C.-H. Robust nonparallel proximal support vector machine with lp-norm regularization. *IEEE Access* **2018**, *6*, 20334–20347.
34. Charrada, A.; Samet, A. Nonlinear Complex LS-SVM for Highly Selective OFDM Channel with Impulse Noise. In Proceedings of the International Conference on Sciences of Electronics, Technologies of Information and Telecommunications, Sousse, Tunisia, 21–24 March 2012; pp. 696–700.
35. Zhou, K.; Zhang, L.; Jiang, M. Enhanced effective SNR prediction for LTE downlink. In Proceedings of the IEEE/CIC International Conference on Communications in China (ICCC), Shenzhen, China, 2–4 November 2015.
36. Study on Narrow-Band Internet of Things (NB-IoT)/enhanced Machine Type Communication (eMTC) Support for Non-Terrestrial Networks (NTN), Document 3GPP TR 36.763 V17.0.0, June 2021. Available online: https://www.tech-in-vite.com/3m36/toc/tinv-3gpp-36-763_a.html (accessed on 18 October 2022).