IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 21, 2021, accepted September 18, 2021, date of publication September 22, 2021, date of current version September 29, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3114619

Extreme Learning Machine (ELM) for Fast User Clustering in Downlink Non-Orthogonal Multiple Access (NOMA) 5G Networks

S. PRABHA KUMARESAN[1], CHEE KEONG TAN[2], (Member, IEEE), AND YIN HOE NG[1]
[1]Faculty of Engineering, Multimedia University, Cyberjaya, Selangor 63100, Malaysia
[2]School of Information Technology, Monash University, Subang Jaya, Selangor 47500, Malaysia

Corresponding author: Chee Keong Tan (tan.cheekeong@monash.edu)

This work was supported by Multimedia University under Grant MMUI/170084.

ABSTRACT Non-orthogonal multiple access (NOMA) has been envisaged as a successor of orthogonal multiple access (OMA) in the fifth generation (5G) networks and beyond because NOMA has been theoretically and empirically proven to be more bandwidth-efficient than OMA. Nevertheless, user clustering (UC) in NOMA is another prevalent issue. To maximize the throughput and fulfill the successive interference cancellation (SIC) constraints, the UC has been formulated as a clustering optimization problem which has been extensively researched in literature. Recently, an artificial neural network-based UC (ANN-UC) scheme has emerged as a viable solution that can optimally cluster users after exhaustive training. However, the ANN model has an extremely slow learning speed, due to the gradient-based back-propagation (BP) algorithm used by the ANN. To address these issues, this paper proposes a novel fast-learning extreme learning machine-based UC (ELM-UC) scheme. Unlike the ANN-UC technique, the input weights and the bias for the hidden layer nodes of ELM are randomly generated and tuning of parameters is not required, thereby leading to a faster learning rate. In this work, the ELM architecture is adapted to operate in NOMA environments where the optimal cluster formation can be predicted rapidly based on the users' channel gains and powers. Performance comparisons with the state-of-the-art UC schemes are investigated via extensive simulations. Remarkably, simulation results demonstrate that the proposed ELM-UC technique can achieve near-optimal performance compared to the brute-force search (B-FS) method and outperforms the existing clustering techniques including ANN-UC and dynamic user clustering (DUC).

INDEX TERMS Extreme learning machine (ELM), non-orthogonal multiple access (NOMA), throughput maximization, user clustering, learning rate.

I. INTRODUCTION

Orthogonal multiple access (OMA) has reached its performance bottleneck which is not capable to accommodate the growing demand of throughput requirement and the system capacity for the future networks, particularly the fifth generation (5G) networks and beyond. Recently, non-orthogonal multiple access (NOMA) has gained significant attention and has become a great paradigm for the new design of multiple access techniques for the 5G systems and beyond 5G systems [1]–[3]. Fundamentally, the emergence of successive interference cancellation (SIC) enables NOMA to support the grouping of multiple users, allowing them to share their radio resources, either in time, frequency or code domain. In a power-domain NOMA system, numerous users are multiplexed by exploiting the channel difference among the users and the composite signals are demultiplexed using the SIC technique that allows the separation of signals from multiple users at the receiver side. It is a common practice that a NOMA system adopts superposition coding (SC) method for signal multiplexing at the sender side and implements the SIC for signal demultiplexing at the receiver side. It has been empirically and theoretically evident that the NOMA scheme can yield higher throughput and capacity as compared to the OMA counterpart, which is one of the key criterion in the 5G networks deployment [4]–[6].

The associate editor coordinating the review of this manuscript and approving it for publication was Diego Oliva.

A. RELATED WORKS

Over the last 5 years, NOMA has been extensively researched and studied, particularly on power allocation and user clustering aspects. In [7], user pairing algorithm for cooperative NOMA transmission has been developed to cluster users with considerably diverse channel conditions. In this work, it has been evidently demonstrated that pairing users in a NOMA cluster provides significant performance gain over OMA. In [8], by utilizing the predetermined grouping of nearby users in one group and distant users in another group, the performance of user pairing of NOMA users was investigated under three different scenarios based on user's distance, i.e., (i) randomly pairing one user from each group; (ii) pairing the nearest near user close to the BS from one group and the nearest far user close to the BS from another group; and (iii) pairing near user similar to scenario (ii) and the farthest far user close to the BS from another group. Based on the investigation, it is proven that user pairing based on scenario (ii) is capable to produce the minimum outage probability and attain the maximum throughput. The work in [9] extensively studies the impact of user pairing on the performance of NOMA-based cognitive radio networks using the fixed power allocation. Similar to the prior work, the results in [9] again manifest the capability of NOMA system that can offer a higher sum rate than the OMA counterpart if an efficient UC algorithm is adopted. Under a fixed-power allocation condition, the strong user is always paired with the weak user based on the significant difference in their channel gains. On the other hand, in cognitive radio assisted NOMA, users that do not experience significant difference in channel gains, i.e., the first strongest user and the second strongest user, are opportunistically paired with each other subject to the SIC constraints.

A user pairing algorithm based on channel state sorting (CSS) is proposed in [10] in which a user with better channel condition is always paired with a user with poorer channel condition to ensure the throughput fairness among the users while improving the overall capacity of the NOMA system. In [11], three different UC schemes for a two-user NOMA network with cell-free massive MIMO are compared: (i) pairing the shortest distance users; (ii) pairing the largest distance users; and (iii) random pairing of users. As compared to the OMA, the sum spectral efficiency of these approaches are significantly improved. Recently, the dynamic UC (DUC) mechanism incorporated with an efficient power allocation under three different scenarios (i.e., 2, 3 and 4 users per cluster) is developed in [12] to maximize the overall system throughput for NOMA systems. The simulation results under these three scenarios indicate that the proposed DUC scheme can achieve efficient and rapid clustering that produces better throughput performance than those in the prior works. Nevertheless, due to the rigid conditions imposed by the proposed DUC (fixed number of users per cluster and pre-determined number of clusters), the capability of the NOMA in terms of bandwidth utilization is undermined. Under such conditions, the users are not allowed to freely form their clusters where in certain scenarios, some users may be forcefully clustered into adverse group, leading to a lower throughput achievement.

To address the limitation of the fixed cluster size and fixed cluster number in the UC optimization in [12], an adaptive UC (AUC) utilizing a Brute-force search (B-FS) has been developed in [13] to thoroughly explore and exploit the diversity of the channel gains of all users to group users collectively. More specifically, this work examines all combinations of UC regardless of the size of the cluster and the number of the clusters formed to find the best cluster that leads to the highest throughput. The conditions on the number of users per cluster and the number of clusters formed in the NOMA are relaxed, which implies that the users can freely form coalition with others or they can form their own singleton cluster (a cluster consisting one user only) or all users can form a grand cluster (a cluster consisting of all users) as long as the system throughput is maximized subject to the SIC constraints. This work has laid down an optimal performance bound of the UC of a NOMA system, which serves as the benchmark of throughput performance for any UC schemes. Nevertheless, since the B-FS approach explores all possible combination to cluster formation, this approach suffers from high computational complexity (the complexity increases exponentially with the number of users) which makes it infeasible for practical implementation. Any variation of the channel gains may activate the UC and the B-FS needs to be executed exhaustively to find the best cluster again. To reduce the computational complexity of the B-FS, particle swarm optimization (PSO) has been employed in optimizing the UC [14], which leads to a sub-optimal sum-throughput. The tradeoff between the throughput performance and complexity is analyzed and it is portrayed that the sum throughput can be sacrificed for lower complexity. It is also revealed that the PSO-based AUC scheme suffers from a local optimality issue where the frequency of obtaining local optimum increases with growing number of users in NOMA systems.

To overcome the scalability issue of PSO-based AUC, artificial neural network (ANN) is adopted in NOMA systems [15] to predict cluster formations after an exhaustive training. The authors make use of the dataset obtained from the B-FS method [13] to train the ANN model with the intention to maximize the overall throughput performance of NOMA system. The implementation of neural network in the UC of NOMA has shed a light on instant UC without the limit on the number of users (unlimited network size), but the training process of the neural network model has become another research issue as a larger network produces a larger dataset which leads to longer training period. Although the ANN-based UC technique is able to provide a satisfactory throughput performance, its training process is time-consuming due to the reliance on the gradient-based back-propagation (BP) learning algorithm to update the weights. From the literature, the application of neural network seems to provide a viable solution to the complicated UC problem, but the arduous training phase of neural network has become the main challenge discouraging the application of neural networks in UC problem.

B. MOTIVATIONS AND CONTRIBUTIONS

Recently, extreme learning machine (ELM) has gained tremendous interest due to its fast learning speed. Unlike the ANN, ELM can learn at an extremely fast manner as it does not require gradient-based BP learning algorithm. More specifically, ELM adopts a single-hidden-layer feedforward neural network (SLFN) which randomly updates the input weights and bias of the nodes at each hidden layer. The least-square method is suggested for ELM to compute the weights linking the nodes from hidden layer to the output layer. Apparently, ELM has demonstrated its effectiveness in training feed-forward neural networks and addressing the shortcomings of the BP algorithm used by neural network for training in [16]–[21]. Extremely fast training could be achieved by ELM as both the weights and biases are randomly assigned while the output weight is analytically determined, thereby eliminating the needs for gradient-based BP learning algorithms to iteratively tune the network parameters. As such, ELM can effectively address the issues that are associated with the gradient-based BP learning schemes, such as slow convergence, local optima, inappropriate learning rate, etc. Furthermore, SLFN is a universal approximator which allows ELM to learn universally by approximating any continuous functions.

Numerous applications have also shown that the ELM model can produce substantially better generalization performance than that of the BP techniques [17], [22]–[24]. Owing to its excellent merits such as exceptionally fast training and high precision, better generalization performance, simple implementation, and universal approximation ability, ELM has been applied in various applications such as ship identification, image quality evaluation, recognition of human action [25]–[27], etc.

Motivated by the aforementioned performance advantages of ELM, an extreme learning machine based user clustering (ELM-UC) technique is proposed in this work to improve the throughput performance of NOMA users at a lower computational complexity and with a higher learning speed. The ELM architecture is adapted to the UC optimization to function as a UC formation predictor based on the useful information at the input. The UC data (i.e., power allocation, channel gains, and cluster formation) collected by using the B-FS method [13] is fed into the proposed architecture to train the ELM model. Since the ELM model can learn at a faster pace compared to the ANN model, more comprehensive data can be generated and fed into the ELM architecture so that the proposed ELM model can predict the clustering formation under any circumstances. Extensive simulations are carried out to investigate the performance of ELM-UC with different setting in various deployment scenarios and the salient features of ELM-UC over the existing UC techniques are discussed in detail.

C. PAPER ORGANIZATION

The remaining sections of the paper are organized as follows. The system model of a downlink NOMA-based 5G cellular network is comprehensively presented in Section II together with the problem statement and throughput formulation. The ELM architecture is briefly discussed in Section III where the novel ELM-UC scheme is developed to optimize UC for a NOMA network. This section also outlines all important features of the proposed ELM-UC schemes as well as the necessary steps for training and testing phases. Simulation results with in-depth analytical discussions are shown in Section IV. Last but not least, the paper ends with some insightful concluding remarks and navigate the readers to some possible future research direction related to this work in Section V.

II. NOMA SYSTEM MODEL

A. NOMA SYSTEM MODEL

A multi-carrier NOMA system with a single base station (BS) that serves Q users using M subcarriers is considered. Let $\mathcal{Q} \triangleq \{1, 2, \ldots, Q\}$ denotes the index set of all users and $\mathcal{M} \triangleq \{1, 2, \ldots, M\}$ be the index set of all subcarriers. Within the cell, the users are randomly and uniformly deployed. The bandwidth of subcarrier m is represented by B_m, for $m \in \mathcal{M}$ and the total bandwidth is denoted by $\mathcal{B} = \sum_{m\in\mathcal{M}} B_m$. It is assumed that the adjacent subcarriers do not suffer from any interference as orthogonal frequency division multiplexing (OFDM) is employed and each subcarrier $m \in \mathcal{M}$ experiences frequency-flat-block-fading at its bandwidth B_m. The channel gain of user q on subcarrier m is represented by g_q^m, for $q \in \mathcal{Q}$, $m \in \mathcal{M}$ and η_q^m is the received noise power of user q on subcarrier m in downlink transmission. Therefore the signal transmitted by the BS to each user q on subcarrier m remains to be active with powers p_q^m when $p_q^m > 0$.

A power-domain NOMA system is considered in which SC is used to multiplex K users on the same subcarrier and the selection of K value depends on the practical constraints of SIC [4]. Let $\mathcal{U}_m \triangleq \{q \in \mathcal{Q} : p_q^m > 0\}$ denotes the set of users multiplexed on subcarrier m. Each subcarrier is configured as a Gaussian broadcast channel with multiple users [28] and SIC is implemented at the receiver side to minimize the intra-band interference. To model the SIC, we must take into account the order of decoding of each user $q \in \mathcal{U}_m$ superimposed on the same subcarrier $m \in \mathcal{M}$. This order of decoding is expressed by a permutation function $\Pi_m : \{1, \ldots, |\mathcal{U}_m|\} \rightarrow \mathcal{U}_m$ where $|\cdot|$ represents the finite set of attribute values. For $r \in \{1, \ldots, |\mathcal{U}_m|\}$, $\Pi_m(r)$ yields the index for the decoded user r. Instead, user $q's$ order of decoding is $\Pi_m^{-1}(q)$. Therefore, the user's signals $\Pi_m(1), \ldots, \Pi_m(r-1)$ are decoded first and removed from the conflicting signal before decoding the signal for $\Pi_m(r)$. In addition, user $\Pi_m(r)$ will be prone to user interference $\Pi_m(s)$, for $s > r$. In contrast, when the previous $|\mathcal{U}_m| - 1$ users are decoded successfully, then $\Pi_m(|\mathcal{U}_m|)$ is decoded last and is not susceptible to any intra-band interferences. Thus, the optimal order of decoding upholds the respective sorting as shown below:

$$\frac{\eta^m_{\Pi_m(1)}}{g^m_{\Pi_m(1)}} \geq \frac{\eta^m_{\Pi_m(2)}}{g^m_{\Pi_m(2)}} \geq \cdots \geq \frac{\eta^m_{\Pi_m(|\mathcal{U}_m|)}}{g^m_{\Pi_m(|\mathcal{U}_m|)}} \tag{1}$$

The communication link is modeled by applying the Shannon capacity formula and the achievable throughput R^m_q of user q on subcarrier m can be expressed as:

$$R^m_q \triangleq B_m \log_2(1 + \frac{g^m_q p^m_q}{\sum_{s=\Pi^{-1}_m(q)+1}^{|\mathcal{U}_m|} g^m_q p^m_{\Pi_m(s)} + \eta^m_q}) \tag{2}$$

According to the SIC order of decoding, user q is only subject to interference from users $\Pi_m(s)$, $s > \Pi^{-1}_m(q)$ The sum system throughput can be further denoted as:

$$R = \sum_{m=1}^{M} \sum_{q=1}^{Q} B_m log_2(1 + \frac{g^m_q p^m_q}{\sum_{s=\Pi^{-1}_m(q)+1}^{|\mathcal{U}_m|} g^m_q p^m_{\Pi_m(s)} + \eta^m_q}) \tag{3}$$

B. CLUSTERING PROBLEM FORMULATION

To successfully implement SIC operation at the receiver, the transmission power at each NOMA user needs to be assigned prudently to ensure the SIC constraints are met. Let's assume that there are three NOMA users sharing a subcarrier m with the channel gains $g^m_1 > g^m_2 > g^m_3$, the transmission power allocated to the BS to the users at User 1 must satisfy the following SIC conditions:

$$g^m_1(p^m_3 - p^m_1 - p^m_2) \geq p_{SIC} \tag{4}$$

$$g^m_1(p^m_2 - p^m_1) \geq p_{SIC} \tag{5}$$

where p_{SIC} is the minimum power gap necessitated by User 1 to distinguish between the decodable and non-decodable signals from a composite signal sent by the BS. The conditions (4) and (5) indicate the power allocation strategy to ensure User 1 to cancel the interference signals caused by Users 2 and 3. These conditions need to be guaranteed for all the allocated subcarriers so that the SIC operations can be carried out successfully at User 1's receiver for all allocated subcarriers. Besides, the transmission power allocation must also satisfy the condition such that $\sum_{m \in M} (p^m_1 + p^m_2 + p^m_3) \leq p_{total}$ where p_{total} is the total available power budget for the BS.

To efficiently cancel the interference signal from User 3 at User 2's receiver by performing SIC, the SIC condition can be expressed as

$$g^m_2(p^m_3 - p^m_1 - p^m_2) \geq p_{SIC} \tag{6}$$

Based on the working principle of NOMA, User 2 is unable to eliminate User 1's interference signal but this interference can be treated as noise because the power allocated to the transmission from the BS to User 1 is smaller compared to that of allocated to User 2 ($p^m_1 < p^m_2$). This phenomenon is also applicable to User 3 who is not capable to eliminate any of the interference signals generated by Users 1 and 2, but the interference is considered as noise because $p^m_1 < p^m_2 < p^m_3$.

Based on conditions (4), (5) and (6), the SIC constraints for a Q-user NOMA system on subcarrier m can be generalized as:

$$g^m_{q-1}(p^m_q - \sum_{i=1}^{q-1} p^m_i) \geq p_{SIC}, q = 2, 3, \ldots, Q, \forall m \in \mathcal{M} \tag{7}$$

Let the user clustering indicator set of a user q be $\theta_q = \{\theta^m_{q,1}\theta^m_{q,2}, \ldots, \theta^m_{q,C}\}, \forall m \in \mathcal{M}$ where $\theta^m_{q,j}$ is a Boolean variable such that $\theta^m_{q,j} = 1$ if a user q is grouped into a cluster j on subcarrier m, otherwise, $\theta^m_{q,j} = 0$. In this cluster formation problem, the total number of cluster is denoted with C, which is a variable (not predetermined) depending on the channel diversity of the NOMA users. Unlike the work in [12] which fixes the total number of clusters in the NOMA-based network, this paper presents a new dynamic clustering problem which flexibly groups users based on their channel heterogeneity and diversity, leading to an unknown value of C. In certain NOMA scenario with high channel diversity, C would be a smaller value, indicating more users are grouped into a cluster to share more subcarriers. On the other hand, if the channel diversity is low, C becomes larger as small channel difference between NOMA users force them to form separate clusters due to incompliance to the SIC conditions. Accordingly, the Q-user clustering indicator set can be represented using the set $\theta = \{\theta_1, \theta_2, \ldots, \theta_Q\}, \forall m \in \mathcal{M}$.

The power allocation strategy of all users on their shared subcarriers within all the cluster can be denoted $P = \{p^m_1, p^m_2, \ldots, p^m_Q\}, \forall m \in \mathcal{M}$. To maximize the throughput of the NOMA system, the joint clustering and power allocation problem for the downlink NOMA system can be formulated as

$$\max_{\theta,P} \sum_{i=1}^{C} \sum_{m=1}^{M} \sum_{q=1}^{Q} \theta^m_{q,i} B_m log_2 \times (1 + \frac{g^m_q p^m_q}{\sum_{s=\Pi^{-1}_m(q)+1}^{|\mathcal{U}_m|} g^m_q p^m_{\Pi_m(s)} + \eta^m_q}) \tag{8}$$

$$\sum_{m=1}^{M} \sum_{q=1}^{Q} p^m_q \leq p_{total} \tag{8a}$$

$$\sum_{m=1}^{M} R^m_q \geq R^{min}_q \tag{8b}$$

$$g^m_{q-1}(p^m_q - \sum_{i=1}^{q-1} p^m_i) \geq p_{SIC}, \quad q = 2, 3, \ldots, \quad Q \forall m \in \mathcal{M} \tag{8c}$$

$$\sum_{i=1}^{C} \sum_{m=1}^{M} \theta^m_{q,i} \leq 1, \forall q \in Q \tag{8d}$$

where constraint (8a) indicates total power budget available at the BS, constraint (8b) ensures the minimum achievable throughput for each NOMA user, constraint (8c) enforces the SIC condition at all users' receivers, and constraint (8d) guarantees that one NOMA user is only assigned to one cluster only. In this context, we assume that subcarriers are preallocated to every user before performing the user clustering and power allocation. Once the users form their clusters, it is assumed that the users will share their subcarriers with other users within the same clusters. The subcarrier allocation is not the main consideration in this work and will be considered in the future research.

The clustering and power allocation problem formulated in (8) is a combinatorial optimization problem which is NPcomplete. Unlike the problem formulated in [12] which can be solved using a dual decomposition method, solving the abovementioned problem is more challenging as the NP-complete issue is aggravated by the relaxation of the rigid constraints on the number of clusters and the number of users that can be grouped in a cluster. Thanks to the work done in [13] which proposes a brute-force search approach, the machine learning starts making inroad into the user clustering problem due to the availability of the clustering dataset. Since the clustering of NOMA users is highly dependent on the channel gains of the users and powers allocated for the transmission, a deep learning method can be utilized to get the model learn about the relationship between the input (channel gains and powers) and output (cluster formation). In this work, Extreme Learning Machine (ELM) approach is adopted due to its fast learning speed and better generalization performance.

III. EXTREME LEARNING MACHINE BASED USER CLUSTERING

In this section, the proposed ELM-UC technique is technically presented. The objective of the proposed methodology is to accurately and rapidly learn the optimal UC of NOMA users based on the input features to maximize the total system throughput utilizing the ELM technique developed in [29]. The proposed ELM structure consists of three layers, known as an input layer followed by a hidden layer, and it ends with an output layer. Unlike the conventional ANN algorithms, which necessitates the iterative adjustment of internal weights and requires the tuning of the training parameters such as the learning rate, etc., ELM-UC randomly assigns the weights connecting the nodes from the input layers to the hidden layers, which does not require BP algorithm. On the other hand, the weights linking the nodes from the hidden layer to the output layer are obtained utilizing the Moore-Penrose (MP) pseudo-inverse according to the least-squares criterion. For these reasons, the training of the ELM can be completed at a much faster rate as compared to other neural networks based on the BP algorithm.

A. DATASET GENERATION

The dataset required for the proposed ELM-UC scheme to learn the optimal formation of clusters which is generated using the B-FS based AUC [13]. The B-FS based AUC scheme proposed in [13] exploits an exhaustive search to explore all the possible clustering combinations to search for the best clustering strategy that is able to yield the highest throughput. Although the B-FS method incurs prohibitive complexity for practical implementation, the method has provided a theoretical performance upper bound and can serve as the data generator to produce different training datasets for the proposed ELM model. More specifically, the dataset consists of a set of users transmit powers and their instantaneous channel gains that corresponds to the best formation of the clusters, leading to the highest throughput. To ensure that the proposed scheme is adaptive and robust against any deployment scenarios, the datasets containing the best cluster formations for different network size with random user positions together with different number of subcarriers in a NOMA system are simulated and collected. The clusters obtained are then numbered in a similar manner as that proposed in [15] for the training purposes.

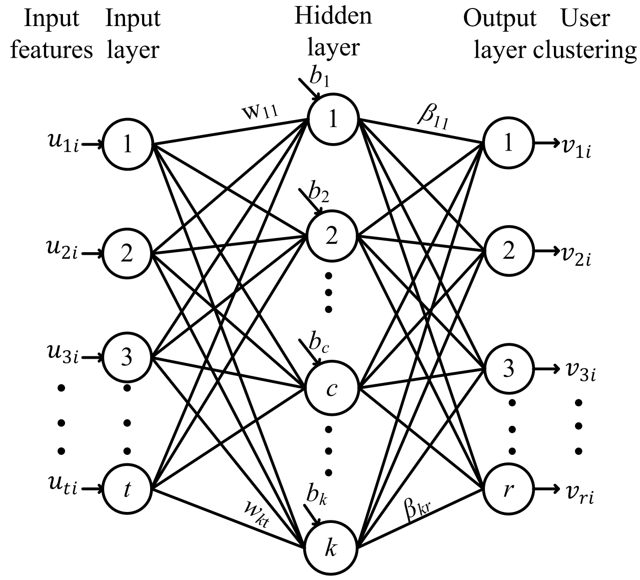

FIGURE 1. Structure of ELM-UC.

B. STRUCTURE OF ELM-UC

The structure of the proposed ELM-UC is depicted in Fig. 1, which consists of t input layer nodes, k hidden layer nodes, and r output layer nodes. The inputs of the proposed ELM-UC is denoted by U and it comprises the channel gains g and initial transmit powers p of Q users. On the other hand, the outputs V corresponds to the clustering information that needs to be obtained at r output layer nodes. As such, for the NOMA system that serves Q users, there are $2Q$ input nodes and Q output nodes, i.e., $t = 2Q$ and $r = Q$. Mathematically, $\{U, V\} = \{u_i, v_i\}$ where $i = 1, 2, \ldots, S$. Thus, the matrix representation of U and V of the training data set are as follows:

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1S} \\ u_{21} & u_{22} & \cdots & u_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ u_{t1} & u_{t2} & \ldots & u_{tS} \end{bmatrix} \tag{9}$$

$$V = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1S} \\ v_{21} & v_{22} & \cdots & v_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ v_{r1} & v_{r2} & \ldots & v_{rS} \end{bmatrix} \tag{10}$$

More explicitly, u_{di} can also be written in terms of the channel gain and transmit power as

$$u_{di} = \begin{cases} g^m_{\left(\frac{d+1}{2}\right)i}, \; for\ d \bmod 2 \neq 0 \\ p^m_{\left(\frac{d}{2}\right)i}, \; for\ d \bmod 2 = 0 \end{cases} \tag{11}$$

where g^m_{ni} and p^m_{ni} signify the channel gain and transmit power of i-th input sample for user n on subcarrier m, respectively.

The input weights and the hidden biases are randomly generated and therefore tuning of parameters in the proposed ELM-UC scheme is not required. On the other hand, the output weights are calculated using MP generalized inverse [30]. Upon defining w_{ij} and β_{ja} as the weight that connects the i-th node of the hidden layer to the j-th node of the input layer and the weight that connects j-th node of the hidden layer to the a-th node of the output layer, respectively, the corresponding matrix representation for input weights W and output weights β can be expressed as:

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1t} \\ w_{21} & w_{22} & \cdots & w_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ w_{k1} & w_{k2} & \cdots & w_{kt} \end{bmatrix} \tag{12}$$

$$\beta = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1r} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{k1} & \beta_{k2} & \ldots & \beta_{kr} \end{bmatrix} \tag{13}$$

The biases for the hidden layer nodes is denoted by $B = [b_1, b_2, \ldots b_t]^T$ and the output layer matrix is represented by $Y = [y_1, y_2, \ldots, y_S]_{r\times S}$, where

$$y_j = \begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{rj} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{k} \beta_{i1} f(w_i u_j + b_i) \\ \sum_{i=1}^{k} \beta_{i2} f(w_i u_j + b_i) \\ \vdots \\ \sum_{i=1}^{k} \beta_{ir} f(w_i u_j + b_i) \end{bmatrix} \tag{14}$$

$f(\cdot)$ denotes the activation function and $j = 1, 2, 3, \ldots, S$. The output layer matrix can be compactly written as

$$H\beta = Y' \tag{15}$$

where H is the output matrix of the hidden layer defined as

$$H = \begin{bmatrix} f(w_1 u_1 + b_1) & f(w_2 u_1 + b_2) & \cdots & f(w_1 u_2 + b_1) \\ f(w_2 u_2 + b_2) & f(w_k u_1 + b_k) & \cdots & f(w_k u_2 + b_k) \\ \vdots & \vdots & \ddots & \vdots \\ f(w_1 u_S + b_1) & f(w_2 u_S + b_2) & \cdots & f(w_k u_S + b_k) \end{bmatrix} \tag{16}$$

C. TRAINING PHASE

The entire data set is divided into training, validation, and testing datasets. In the training phase, the desired user clustering is known to the ELM-UC, i.e. $y_j = v_i$. The proposed ELM-UC is trained to determine the least-squares solution β for $min_\beta \left\| H\beta - Y' \right\|$ and the smallest norm least- squares solution can be written as

$$\beta = H^\dagger Y' \tag{17}$$

where $H^\dagger$ is the MP generalized inverse of H.

Generally, the k has a significant impact on training the neural network. If the number of hidden nodes is the same as the number of training samples, i.e. $k = S$, the mean squared error (MSE) of the conventional SLFNs can approach zero [29]. For the proposed ELM-UC, the number of hidden nodes is significantly less than the number of training samples, i.e. $k \ll S$, which causes the MSE to approach an arbitrary value $\xi > 0$,

$$MSE = \frac{1}{S} \sum_{i=1}^{S} (u_i - y_i)^2 < \xi \tag{18}$$

According to ridge regression theory, the stability and generalization ability of ELM could be enhanced by including a positive regularization term $1/\lambda$ in the calculation of β as follows:

$$\beta = (\frac{I}{\lambda} + H^T H)^{-1} Y' \, for \; < S \tag{19}$$

where I is the identity matrix.

D. VALIDATION AND TESTING PHASES

Once the training is completed, the proposed model is validated by using the validation data samples. The purpose for validation is to tune the hyper-parameters, such as the value of k and the value of the regulation parameter, so that the machine learning model is able to make accurate prediction. In the testing phase, the input features of testing data samples are fed into the model to predict the UC of NOMA users. Based on the predicted clustering formation, the transmit power is equally allocated to all users for all the clusters. For simplicity, the equal power allocation method is chosen in this context because power allocation problem is not the main focus of this work. For a fair comparison, other benchmarking techniques will also adopt the similar equal power allocation method.

The performance of ELM-UC can be evaluated in terms of throughput using (3) based on the predicted clustering formation and allocated power. The accuracy of the prediction is denoted as follows:

$$\text{Accuracy} = \frac{\mathcal{N}_{\mathbb{C}}}{\mathcal{N}_S} \tag{20}$$

where $\mathcal{N}_{\mathbb{C}}$ symbolizes the number of correctly predicted samples and $\mathcal{N}_S$ represents the total number of samples.

Algorithm 1 summarizes the operational steps of the proposed ELM-UC technique.

E. SALIENT FEATURES OF ELM-UC

Since the underlying architecture of the proposed UC scheme is ELM, ELM-UC inherits all the inherent features of ELM which can be summarized as follows:

Algorithm 1 Proposed ELM-UC Technique

Training phase

Step 1: Randomly generate W and B

Step 2: Calculate H for the training dataset using (16).

Step 3: Compute β for the training dataset using (17).

Validation phase

Step 1: Calculate H for the validation dataset using (16).

Step 2: Predict the UC for the validation dataset using (15).

Step 3: Compute the throughput for the validation dataset using (3).

Step 4: Repeat the training phase and steps 1-2 of the validation phase for different configurations of ELM-UC, i.e. different types of activation functions, different settings of k and λ.

Step 5: Select the best setting of ELM-UC which corresponds to the highest throughput performance.

Testing phase

Step 1: Calculate H for the testing dataset using (16) based on the best setting of ELM-UC obtained in the step 5 of validation phase.

Step 2: Predict the UC for the testing dataset using (15).

Step 3: Calculate the MSE for the testing dataset using (18).

Step 4: Compute the accuracy for the testing dataset using (20).

Step 5: Determine the throughput for the testing dataset using (3).

1) Extremely fast training could be achieved by ELM-UC as both W and B of ELM-UC are randomly assigned while β is analytically determined using (16), thereby eliminating the needs for gradient-based BP learning algorithms to iteratively tune the network parameters. Consequently, ELM-UC can effectively overcome the shortcomings of gradient-based BP learning approaches, such as slow convergence, local optima, improper learning rate, etc.
2) Despite the fact that SLFN is a universal approximator, most of its learning algorithms do not fulfill the universal approximation property [32]. On the other hand, as proven in [18], ELM is a universal learner which is capable of approximating any continuous functions.
3) Since (10) attempts to seek for the smallest norm of β among the least-squares solutions, the solution for β is unique and good generalization performance could be attained by ELM-UC. More explicitly, as pointed out by Barlett's theory, the generalization performance of SLFN tends to improves when both the training errors and the norm of the weights reduce [31].

IV. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

To show the performance of the proposed ELM-UC, extensive simulation is conducted for different NOMA deployments. Since the hyper-parameters for the ELM-UC critically impacts the performance of the ELM-UC, a comprehensive simulation is performed to obtain the optimal setting of the ELM-UC via the training phase. In this extensive simulation, various activation functions will be explored together with the number of hidden layers, number of nodes for each hidden layer, regulation parameters, and length of training data samples. This simulation will conclude the best hyper-parameters for the ELM-UC so that the best ELM-UC approach can be simulated and compared fairly with other UC schemes. More specifically, the throughput achievement of the proposed technique is benchmarked against those of the conventional OMA (lower-bound performance), DUC, B-FS based UC (upper-bound performance), and ANN-UC. The simulation setting for the proposed ELM-UC scheme is summarized in Table 1.

TABLE 1. Simulation setting.

Parameter	Value
Number of NOMA users	12
Number of input layer nodes	24
Number of hidden layer nodes	10, 20, 30, 40, 50, 60
Number of output layer nodes	12
Regulation parameter	1000, 2000, 3000, 4000, 5000, 6000
Size of data	12000
Length of training data	6200, 7200, 8400
Length of validation data	1800
Length of testing data	1800

First, the performance of the training accuracy for different activation functions used by the proposed ELM-UC is compared in Fig. 2. Fig. 2 shows the training accuracy (%) of ReLu, Sigmoid, Sine, and Tanh functions used to activate the nodes in the proposed ELM-UC with respect to different k. The proposed ELM model is trained with 8400 training samples and regulation parameter is fixed at 5000. Choosing the most suitable activation function is crucial as the activation function has a significant on the performance of the proposed scheme. In Fig. 2, it can be seen that ReLu function consistently outperforms other activation functions in terms of accuracy for different k considered. It is also worth-mentioning that ReLu function achieve linear improvement in terms of training accuracy when the k value is increased. On the other hand, the training accuracy of ELM-UC using Tanh, Sine, and Sigmoid activation functions can merely achieve 54%- 60% of accuracy. Furthermore, the ELM-UC scheme adopting Tanh, Sine and Sigmoid functions does not show any significant improvement as the k value increases. Undoubtedly, it is always recommended to employ ReLu as the activation function for the proposed ELM-UC model which can achieve 84% of training accuracy when $k = 50$.

Next, apart from the training accuracy, the investigation is also carried out to analyze the training time required by the aforementioned 4 activation functions during the training phase. In Fig. 3, the training time and training accuracy of the ELM-UC scheme using the ReLu, Sigmoid, Sine, and Tanh function are compared in an ELM architecture with $k = 50$

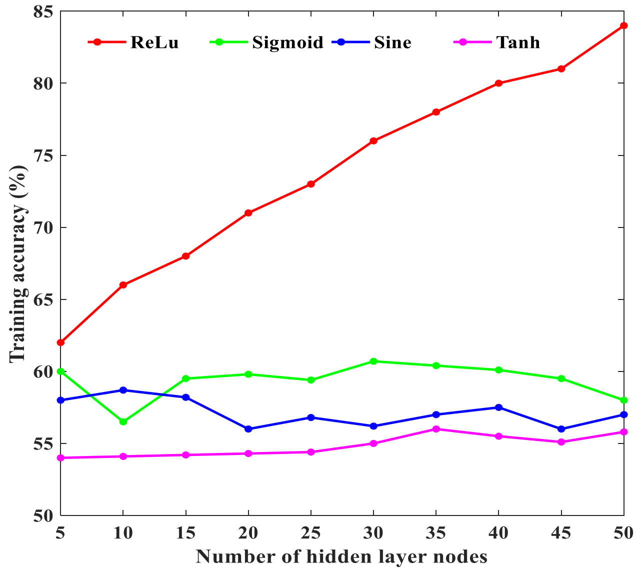

FIGURE 2. Accuracy of the ELM-UC for different numbers of hidden layer nodes and activation functions in training phase.

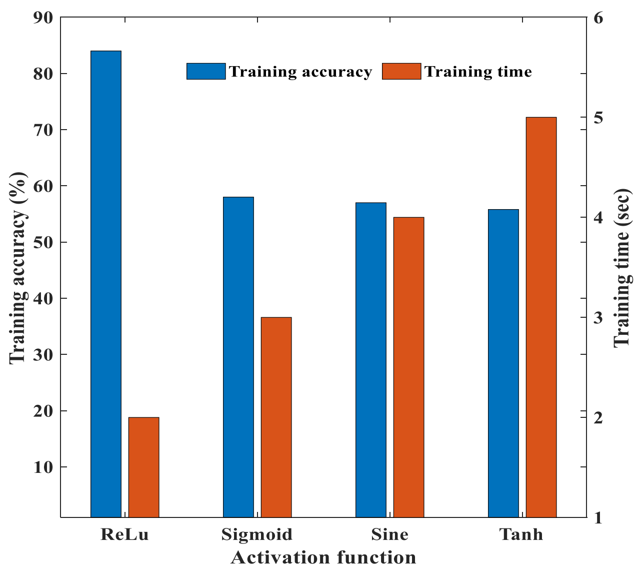

FIGURE 3. Training accuracy and training time of ELM-UC for different types of activation functions.

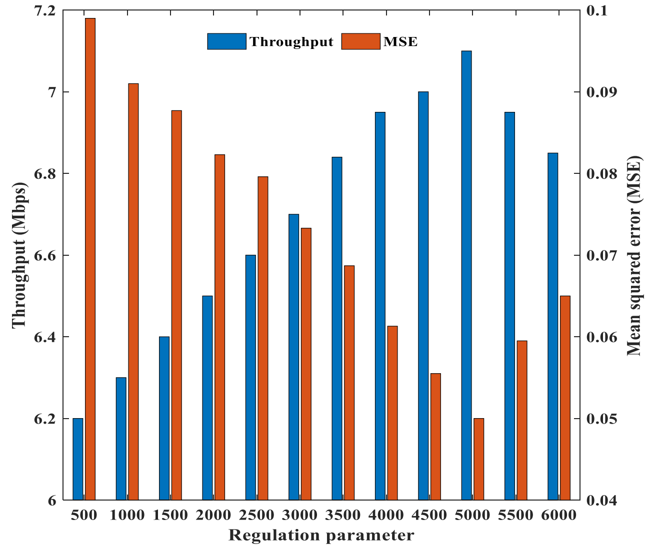

FIGURE 4. Throughput and MSE performance of ELM-UC during the training phase for different setting of regulation parameter.

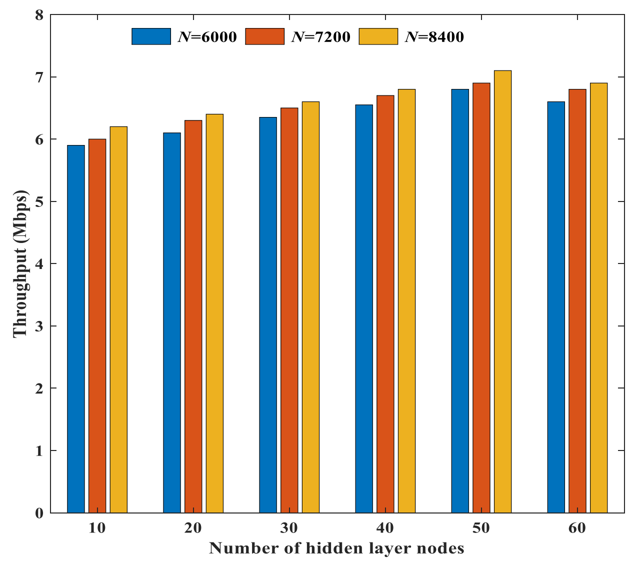

FIGURE 5. Throughput performance of ELM-UC for different numbers of hidden layer nodes and different lengths of training data samples N in testing phase.

and its regulation parameter is adjusted to be 5000. It is noticed in Fig. 3 that the ReLu function enables the ELM-UC to be trained within 20 seconds, but the training time required by the Sigmoid, Sine and Tanh functions is approximately 2-4 times more than that required by ReLu. Again, it is evidently demonstrated that ELM-UC with ReLu could achieve much better training accuracy with much shorter training time as compared to other activation functions. Hence, for the following simulations, the ReLu function will always be used as the activation function in the proposed ELM-UC scheme.

Fig. 4 displays the throughput performance and MSE of the proposed ELM-UC for different settings of regulation parameter with $k = 50$ and 8400 training samples. It is noted that the throughput performance improves and the MSE reduces in a linear scale as the regulation parameter is increased from 500 to 5000. However, as the regulation parameter is increased beyond 5000, the opposite trend is observed. This is attributed to the fact that the regulation parameter plays an essential role in introducing the trade-off between the MSE minimization and the norm of output weights. From the Fig. 4, the optimal regulation parameter is 5000 as it yields the best throughput and MSE performances.

Fig. 5 illustrates the throughput attainments of the proposed ELM-UC scheme for different lengths of training samples N and different values of k in testing phase. The regulation parameter is set to the optimal value 5000 and

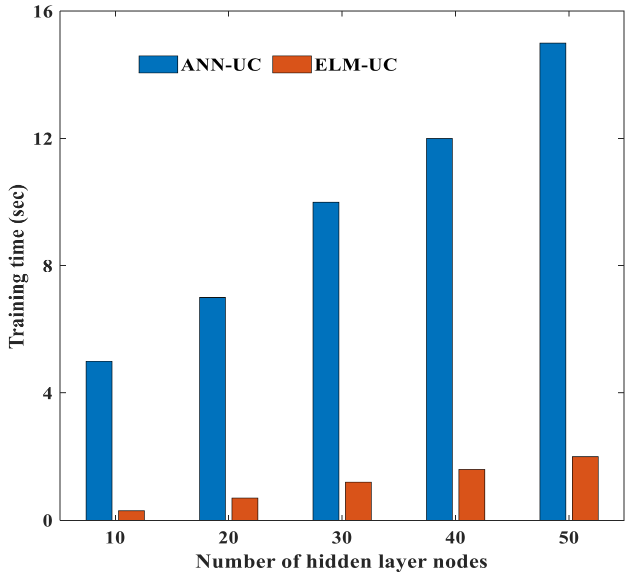

FIGURE 6. Training time of various machine learning based UC techniques for different numbers of hidden layer nodes.

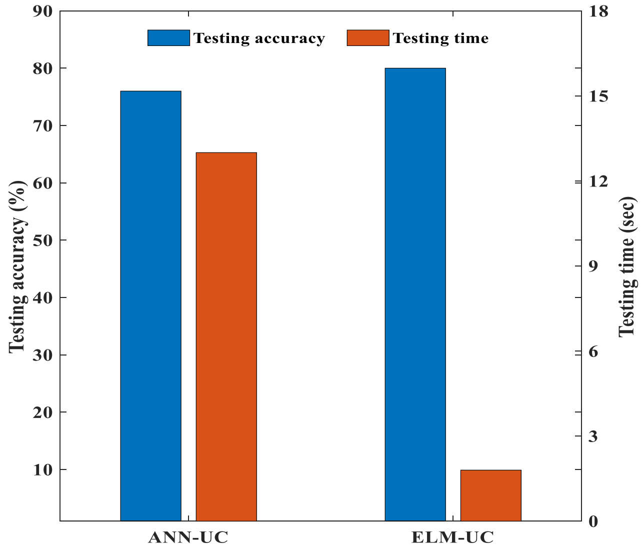

FIGURE 7. Accuracy and time for various machine learning based UC techniques during the testing phase.

ReLu is suggested as the activation function. As expected, it is observed that the throughput performance improves as the length of the training samples increases. Besides that, it is also observed that k value has a significant impact on the throughput performance. More precisely, the proposed ELM-UC scheme exhibits better throughput performance as k is increased from 10 to 50 owing to the improved modeling capability. However, the throughput performance starts to deteriorates when the value of k is further increased to 60 due to the over fitting of the ELM model. As such, we can conclude that the optimal value of k for the scenario considered is found to be 50.

Fig. 6 depicts the training time of two machine learning based UC schemes for different k. In general, it is observed that the training time is proportional to k. As expected, the training time for ANN-UC is significantly longer than that of the ELM-UC. Specifically, for the case of $k = 50$, ELM-UC requires less than 5 seconds for training but ANN-UC takes more than 10 seconds, doubling the time required by the former. This is due to the fact that the input and output weights of ANN-UC are iteratively updated by the BP algorithm. On the other hand, for the case of ELM-UC, the input weights are randomly assigned and the output weights are analytically determined. With this beneficial feature, the ELM-UC is considered to be more feasible when the UC is required to adapt to a vast variety of deployment scenarios when there is a huge volume of training data, which can be rapidly learned by the proposed ELM-UC compared to the ANN-UC.

Fig. 7 compares the accuracy and time required for ANN-UC and ELM-UC during the testing phase. As compared to ANN-UC, ELM-UC can achieve better accuracy with substantial reduction in testing time as it utilizes the smallest norm of output weights to predict the UC and its learning mechanism is performed without iteratively tuning the hidden nodes. From Fig. 7, it is again proven that ELM is a better machine learning model compared to ANN.

Fig. 8 presents the MSE and throughput performance of ANN-UC and ELM-UC for different k values with 1800 testing samples and ReLu activation function during the testing phase. From Fig. 8(a), it is observed that the MSE of ANN-UC reduces as k is increased from 5 to 20. However, as the value of k is increased beyond 20, the MSE of ANN-UC increases accordingly. This is attributed to the fact that ANN-UC with $k = 24$ are sufficient for learning the UC of NOMA users and further increasing k value will lead to overfitting which results in a larger MSE. Owing to the superior generalization capability of ELM-UC, the MSE of ELM-UC is smaller as compared to that of the ANN-UC for the different k values considered. As expected, the MSE achievement substantially influences the throughput performance of the techniques as shown in Fig. 8(b). More specifically, the throughput performance improves as the MSE increases. From Fig. 8(b), it can also be seen that the throughput performance of ELM-UC outperforms that of the ANN-UC and the performance improvement gap becomes wider when a larger value of k is used.

Fig. 9 shows the throughput comparison of OMA with other UC techniques to investigate how each technique scales with the number of users in the testing phase. It is demonstrated that the increasing rate of throughput performance using NOMA UC techniques is higher than that of OMA. This is mainly because the NOMA system allows each subcarrier to be shared among the users within the same cluster, leading to more frequency resources for all users under a circumstance where the co-channel interference is under control (the implementation of SIC can eliminate the interfer-

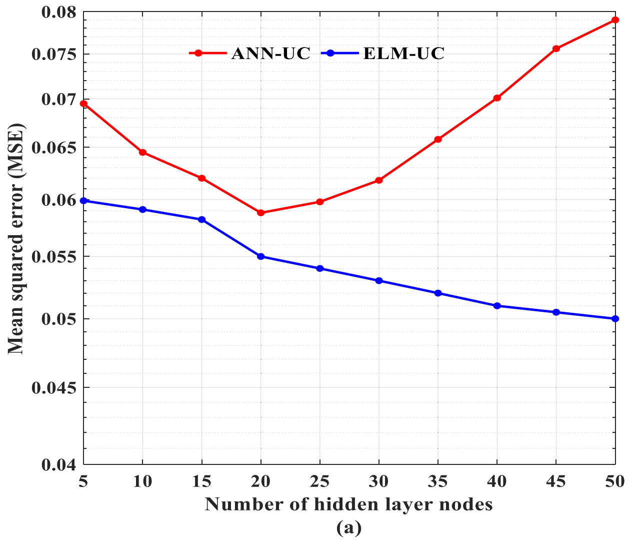

FIGURE 8. MSE and throughput performance of various machine learning based UC techniques for different numbers of hidden layer nodes during the testing phase.

ence effectively for the users with stronger channel gains). In comparison with the optimal B-FS-UC, the performance of the proposed ELM-UC is only slightly inferior with a maximum performance degradation of 7.40%. Besides that, it is also noteworthy that the proposed ELM-UC yields superior throughput performance for all the cases considered as compared to those of OMA, DUC, and ANN-UC. More explicitly, for all the different number of users considered, the proposed ELM-UC is found to exhibit throughput performance gains of 4.41% to 22.72% over the ANN-UC counterpart. This can be explained as follows. Generally, ANN-UC is prone to converge to local minima as it employs gradient-based BP algorithm to update its weights [15]. Consequently, the superiority of ELM-UC comes into play as its output weights are analytically computed based on the minimum norm solution, which results in improved generalization capability as

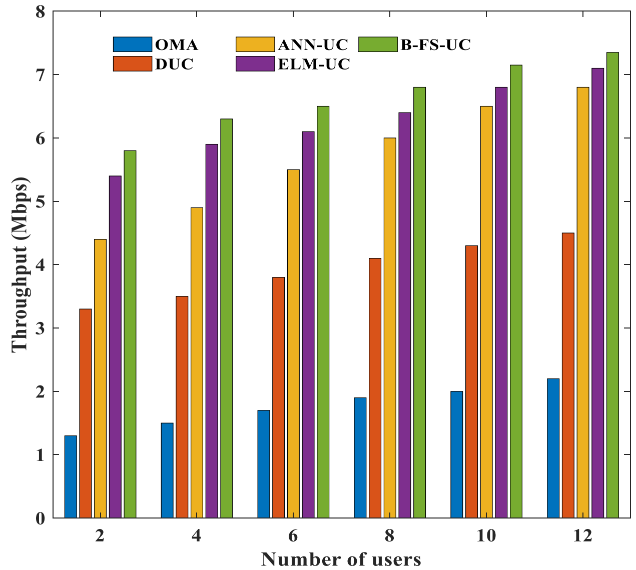

FIGURE 9. Throughput performance of OMA and various UC techniques for different number of users during the testing phase.

reported by the Barlett's theory [31] and the issue of local minima could also be avoided [17].

V. CONCLUSION

This paper proposed a novel fast-learning and low-complexity UC technique to effectively group the users in NOMA systems using an extreme learning machine approach. The proposed ELM-UC scheme is developed to learn the non-linear relationship between the input features and output of UC via a training dataset with tuning-free learning strategy. The relatively long learning time issue which plagues neural network-based UC schemes that adopt BP learning algorithm can be effectively addressed by ELM-UC as its output weights are solved in a single step without the need of time-consuming BP algorithm. Comprehensive simulation results reveal that the proposed technique could attain substantial improvement in terms of throughput and MSE over the existing schemes with a significantly lower complexity and shorter execution time. Even though the proposed number of hidden neurons in the ELM-UC is still acceptable, it manifests an increasing trend if we intend to further increase the throughput performance. Therefore, to optimize the ELM-UC architecture for the best throughput performance, the hidden neuron pruning can be investigated in the future work.

REFERENCES

[1] A. Benjebbour, Y. Saito, Y. Kishiyama, A. Li, A. Harada, and T. Nakamura, "Concept and practical considerations of non-orthogonal multiple access (NOMA) for future radio access," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst. (ISPACS)*, Nov. 2013, pp. 770–774.

[2] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink non-orthogonal multiple access (NOMA)," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 611–615.

[3] A. Benjebbour, A. Li, Y. Saito, Y. Kishiyama, A. Harada, and T. Nakamura, ''System-level performance of downlink NOMA for future LTE enhancements,'' in *Proc. IEEE Global Commun. Conf. (GC Wkshps)*, Dec. 2013, pp. 66–70.
[4] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[5] G. Caire and S. Shamai (Shitz), ''On the achievable throughput of a multiantenna Gaussian broadcast channel,'' *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
[6] K. Higuchi and A. Benjebbour, ''Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access,'' *IEICE Trans. Commun.*, vol. 98, no. 3, pp. 403–414, Mar. 2015.
[7] Z. Ding, M. Peng, and H. V. Poor, ''Cooperative non-orthogonal multiple access in 5G systems,'' *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug. 2015.
[8] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, ''Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer,'' *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 938–953, Apr. 2016.
[9] Z. Ding, P. Fan, and H. V. Poor, ''Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions,'' *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Sep. 2015.
[10] H. Zhang, D. K. Zhang, W. X. Meng, and C. Li, ''User pairing algorithm with SIC in non-orthogonal multiple access system,'' in *Proc. Int. Conf. Commun.*, May 2016, pp. 1–6.
[11] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, L. Hanzo, and P. Xiao, ''On the performance of cell-free massive MIMO relying on adaptive NOMA/OMA mode-switching,'' *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 792–810, Feb. 2020.
[12] M. S. Ali, H. Tabassum, and E. Hossain, ''Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems,'' *IEEE Access*, vol. 4, pp. 6325–6343, 2016.
[13] S. P. Kumaresan, T. C. Keong, L. C. Kwang, and N. Y. Hoe, ''Adaptive user clustering for downlink nonorthogonal multiple access based 5G systems using brute-force search,'' *Trans. Emerg. Telecommun. Technol.*, vol. 31, no. 11, p. e4098, Aug. 2020.
[14] Y. H. Ng, C. K. Lee, S. P. Kumaresan, and C. K. Tan, ''Low-complexity particle swarm optimization based adaptive user clustering for downlink non-orthogonal multiple access deployed for 5G systems,'' *World Rev. Sci., Technol. Sustain. Develop.*, vol. 1, no. 1, p. 1, 2020.
[15] S. P. Kumaresan, C. K. Tan, and Y. H. Ng, ''Efficient user clustering using a low-complexity artificial neural network (ANN) for 5G NOMA systems,'' *IEEE Access*, vol. 8, pp. 179307–179316, 2020.
[16] G. B. Huang and C. K. Slew, ''Extreme learning machine: RBF network case,'' in *Proc. 8th Int. Conf. Control, Automat., Robot. Vis. (ICARCV)*, vol. 2, Dec. 2004, pp. 1029–1036.
[17] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, ''Extreme learning machine: Theory and applications,'' *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
[18] G.-B. Huang, L. Chen, and C.-K. Siew, ''Universal approximation using incremental constructive feedforward networks with random hidden nodes,'' *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 879–892, Jul. 2006.
[19] G. Feng, G.-B. Huang, Q. Lin, and R. Gay, ''Error minimized extreme learning machine with growth of hidden nodes and incremental learning,'' *IEEE Trans. Neural Netw.*, vol. 20, no. 8, pp. 1352–1357, Aug. 2009.
[20] R. Zhang, Y. Lan, G.-B. Huang, Z.-B. Xu, and Y. C. Soh, ''Dynamic extreme learning machine and its approximation capability,'' *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 2054–2065, Dec. 2013.
[21] M.-B. Li, G.-B. Huang, P. Saratchandran, and N. Sundararajan, ''Fully complex extreme learning machine,'' *Neurocomputing*, vol. 68, nos. 1–4, pp. 306–314, Oct. 2005.
[22] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, ''Extreme learning machine for regression and multiclass classification,'' *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.
[23] N. Liu and H. Wang, ''Ensemble based extreme learning machine,'' *IEEE Signal Process. Lett.*, vol. 17, no. 8, pp. 754–757, Aug. 2010.
[24] J. Cao, Z. Lin, G.-B. Huang, and N. Liu, ''Voting based extreme learning machine,'' *Inf. Sci.*, vol. 185, no. 1, pp. 66–77, Feb. 2012.
[25] J. Tang, C. Deng, G. B. Huang, and B. Zhao, ''Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine,'' *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1174–1185, Mar. 2015.
[26] S. Suresh, R. V. Babu, and H. J. Kim, ''No-reference image quality assessment using modified extreme learning machine classifier,'' *Appl. Soft Comput.*, vol. 9, no. 2, pp. 541–552, Mar. 2009.
[27] R. Minhas, A. Baradarani, S. Seifzadeh, and Q. M. J. Wu, ''Human action recognition using extreme learning machine based on visual vocabularies,'' *Neurocomputing*, vol. 73, nos. 10–12, pp. 1906–1917, Jun. 2010.
[28] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.
[29] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, ''Extreme learning machine: A new learning scheme of feedforward neural networks,'' in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Aug. 2004, pp. 985–990.
[30] Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, and G.-B. Huang, ''Evolutionary extreme learning machine,'' *Pattern Recognit.*, vol. 38, no. 10, pp. 1759–1763, Oct. 2005.
[31] P. L. Bartlett, ''The sample complexity of pattern classification with neural networks: The size of the weights is more important than the size of the network,'' *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 525–536, Mar. 1998.
[32] G. Huang, G.-B. Huang, S. Song, and K. You, ''Trends in extreme learning machines: A review,'' *Neural Netw.*, vol. 61, pp. 32–48, Jan. 2015.

S. PRABHA KUMARESAN received the B.E. degree in electronics and communication engineering from IFET College of Engineering, India, in 2004, and the M.Tech. degree in advanced communication systems from SASTRA University, India, in 2013. She is currently pursuing the Ph.D. degree with the Faculty of Engineering, Multimedia University, Malaysia, under the supervision of Dr. Tan Chee Keong and Dr. Ng Yin Hoe. She worked as a Lecturer at the Engineering College, from 2004 to 2009 and from 2014 to 2015. Her research interests include NOMA, 5G, machine learning, and optimization techniques.

CHEE KEONG TAN (Member, IEEE) received the B.Eng. degree in electronics (telecommunication), the M.Eng.Sc. degree in information, communication, and technology, and the Ph.D. degree in information, communication, and technology from Multimedia University, Malaysia. He is currently a Senior Lecturer with the School of Information Technology, Monash University, Malaysia. He has carried out projects for telecommunication companies and cellular service providers, which led to the development of a few patents on wireless algorithm and protocol design. He is the main contributing author to more than 20 international journal articles. His current research interests include radio resource management, 5G networks, indoor positioning scheme, game theory, machine learning, and artificial intelligence.

YIN HOE NG received the B.Eng. degree (Hons.) in electronics engineering and the Master of Engineering Science and Ph.D. degrees from Multimedia University, in 2004, 2008, and 2013, respectively. He is currently a Senior Lecturer with the Faculty of Engineering, Multimedia University. His current research interests include advanced signal processing techniques for digital communication systems, machine learning, and indoor positioning.