2018 IEEE 4th International Conference on Computer and Communications

DGPA-Based Tone Reservation for PAPR Reduction in OFDM Systems

Jun Li*, Renjie Zhang*, Yajun Wang†, Tingting Liu‡, Ran Ding*
*School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, China
†Department of Information and Computational Sciences, JiangSu University of Science and Technology, Zhangjiagang, China
‡School of Communication Engineering, Nanjing Institute of Technology, Nanjing, China
e-mail: {jun.li,renjie.zhang,117104010631}@njust.edu.cn, wangyj1859@just.edu.cn, liutt@njit.edu.cn

Abstract—**In this paper, we concentrate on the high peak-to-average power ratio (PAPR) reduction issue in the orthogonal frequency division multiplexing (OFDM) system based on the tone reservation (TR) method. We first apply the discrete generalized pursuit algorithm (DGPA) to search the optimal peak reduction tone (PRT) set. Then, we compare the performance and computational complexity (CC) between the DGPA and other techniques in finding the PRT set. Furthermore, we apply the PRT set searched by the DGPA in the TR method to reduce the high PAPR. And we compare the PAPR reduction and the bit error rate (BER) performance between three methods. Finally, the simulation results verify that the proposed method in this paper can achieve the similar BER performance as the other benchmarks but have better performance in PAPR reduction.**

Keywords-OFDM; peak-to-average power ratio; tone reservation; discrete generalized pursuit algorithm

I. INTRODUCTION

As one of the multi-carrier modulation techniques, the orthogonal frequency division multiplexing (OFDM) is widely used in the wireless technology. OFDM has plenty of advantages, such as offering a considerable high spectral efficiency, providing good resistance to multi-path fading and avoiding the in-band distortion and out-of-band radiation. However, there are still some challenging disadvantages in the OFDM systems. One disadvantage is that the OFDM system is easily affected by the spectral deviation so that this problem may influence the orthogonality of the sub-carriers in the whole OFDM systems. The other disadvantage is that the OFDM systems have a potential occurrence of the high peak-to-average power ratio (PAPR) of the transmitted signals.

Due to the above disadvantages, the high PAPR should be reduced in the OFDM systems. Until now, many works have been carried out to reduce high PAPR. The overview of PAPR reduction [1] shows us various approaches, including clipping and filtering [2], coding schemes [3], selective mapping (SLM) [4], partial transmitted sequence (PTS) [5], tone injection (TI) [6], tone reservation (TR) [7] and active constellation extension (ACE) [8] etc. These methods can be divided into two main categories that one is the signal distortion approaches, such as clipping and filtering, and the other one is the signal non-distortion schemes like TR and coding.

Among these high PAPR reduction methods, the TR technique proposed by Tellado [9] is an efficient way to reduce PAPR without the transmission of side information and the distortion of the transmitted signals. The key point of TR method is to select the appropriate peak reduction tone (PRT) set. However, it is a nondeterministic polynomial-time (NP)-hard problem to search the optimal PRT set. Consequently, the sub-optimal PRT set is usually taken into consideration. The paper [7] applies the genetic algorithm to search the PRT set. The authors in [10] use the cross entropy to find the PRT set. But these intelligent algorithms converge a little slowly and have high computational complexity (CC). In order to get the PAPR reduced signal, an adaptive scaling (AS) algorithm [11] is proposed to clip the transmitted signal at the preset threshold based on the TR. However, the AS method is difficult to choose the value of the threshold to clip the signal. Then, the adaptive amplitude clipping (AAC) algorithm [7] is put forward to determine the clipping threshold based on the TR. But AAC method has a high CC which needs to be decreased.

The discrete generalized pursuit algorithm (DGPA) is evolved from the reinforcement learning which is widely used in the learning automata. The DGPA is proposed [12] to pursuit all the actions with higher reward estimations than the current selected action, which minimizes the probability of pursuing a wrong action than conventional schemes. Besides, the DGPA has a faster converge rate than old algorithms.

The DGPA has been applied to various aspects. The papers [13], [14] develop a socially aware distributed caching strategy based on the DGPA to optimize the cache placement operation in the device to device (D2D) communication networks. The paper [15] applies the DGPA to optimize the content caching placement in order to minimize the downloading latency in the heterogeneous cellular networks. The authors [16] take advantage of the DGPA to optimize the content placement problems in the caching theory under the physical and social constraints in order to maximize the throughput of the whole D2D communications.

This paper is the first research which applies the DGPA to search the PRT set for PAPR reduction in OFDM systems. And some lights are shed as follows.

- First, we apply the DGPA to search the PRT set for PAPR reduction in the TR-based OFDM systems.

- Then, we develop a new method based on the DGPA to reduce PAPR and decrease the CC.
- Next, we compare the CC and PAPR reduction performance between this method and other benchmarks.
- Finally, the simulation results validate the efficiency of our proposed method and show the superiority of the method.

The notations in this paper are listed as follows. $(.)^T$ means the transpose of a matrix or a vector. E[.] denotes the expectation. $||.||_\infty$ means the infinite norm of a vector.

II. OFDM Systems

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc.

In the OFDM systems, the randomly generated bit data streams are encoded and then transferred from the serial to parallel. N independent OFDM data symbols X_k are modulated by the phase shift keying (PSK) or quadrature amplitude modulation (QAM). One OFDM symbol consists of N subcarriers whose frequency gap is the same. The OFDM block is defined as $X = [X_0, X_1,\dots, X_{N-1}]^T$. The block diagram of the transmitter in the OFDM systems is shown in the Fig. 1.

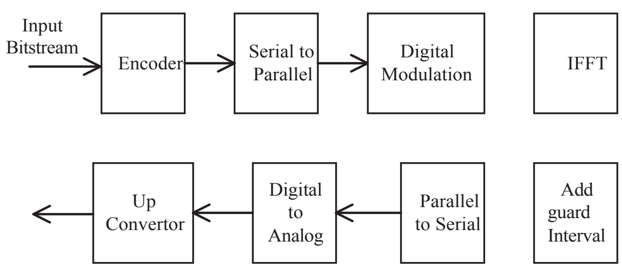

Figure 1. The block diagram of the transmitter in the OFDM system.

When the OFDM signal is over-sampling at the value of L, the discrete time baseband OFDM signal with N sub-carriers after the inverse fast fourier transform (IFFT) is generated as

$$x_n = \frac{1}{\sqrt{LN}} \sum_{k=0}^{N-1} X_k e^{\frac{j2\pi nk}{LN}}, n = 0,1,\dots, LN-1. \quad (1)$$

Here, L is the oversampling factor. x_n is the signal in the time domain and X_k is the k-th frequency signal in the frequency domain. The over-sampling is to do zero-padding operation in the OFDM symbol by adding (L-1)* N zeros.

A. PAPR and CCDF

Then, the PAPR of this OFDM time signal x is defined as

$$PAPR(x) = \frac{\max\limits_{0 \le n < LN} [|x_n|^2]}{E[|x_n|^2]}, \quad (2)$$

where $\max[|x_n|^2]$ means the instantaneous maximum power of the time signal x and the $E[|x_n|^2]$ denotes the average power of the time signal x.

As one of characteristics of the PAPR, the distribution of PAPR, which bears stochastic characteristics in OFDM systems, can be expressed in terms of complementary cumulative distribution function (CCDF) [1]. The CCDF is used to measure the performance of the PAPR reduction. The CCDF means the probability that the PAPR of one OFDM symbol exceeds the preset threshold $PAPR_0$, which is defined as

$$CCDF = \Pr(PAPR > PAPR_0). \quad (3)$$

B. Tone Reservation

In the TR scheme, parts of the sub-carriers named PRT set generate the peak reduced signal to reduce PAPR. These peak reduced tones do not carry any data information. Here, we define that M tones are selected as the PRT set for PAPR reduction. Therefore, the remaining N-M tones are for data transmission. M is the length of PRT set and M < N. The peak reduced signal $c = [c_0, c_1,\dots, c_{N-1}]^T$ is generated by the PRT set. Then after adding the peak reduced signal c with the original time signal x, the combined time signal a is defined as

$$a = x + c = Q(X + C) \quad (4)$$

where Q means the IFFT matrix. The X and C denote the original frequency signal and peak reduced frequency signal respectively. Besides, the X and C are orthogonal to each other in the frequency domain.

$$X(k) + C(k) = \begin{cases} X(k), k \in R^c, \\ C(k), k \in R, \end{cases} \quad (5)$$

where R means the index set of the reserved tones and R^c denotes its complementary set.

Based on (4), the PAPR of the combined time signal a is defined as

$$PAPR(a) = \frac{\max\limits_{0 \le n < N} E[|x_n + c_n|^2]}{E[|x_n|^2]}, \quad (6)$$

where c should be chosen optimally in order to minimize the peak amplitude of the combined time signal a.

The frequency domain kernel $P = [P_0, P_1,\dots, P_{N-1}]^T$ related with the PRT set is defined as

$$P_n = \begin{cases} 0, n \in R^C \\ 1, n \in R \end{cases}, n = 0,1,\dots, N-1 \quad (7)$$

Then the time domain kernel $p = [p_0, p_{,\dots,} p_{N-1}]^T$ can be acquired by

$$p = QP. \quad (8)$$

Usually, in the TR-based scheme, we use the secondary peak of the time domain kernel p to measure the performance for the selection of the PRT set. The secondary peak of the time signal kernel p is defined by

$$f(p) = ||[p_1, p_2,\dots, p_{N-1}]^T||_\infty. \quad (9)$$

In order to acquire the optimal PRT set, the optimization problem is formulated as

$$R^* = \arg\min_R \| [p_1, p_2, .., p_{N-1}]^T \|_\infty . \quad (10)$$

The performance for PAPR reduction in the TR method in OFDM system depends on the selection of the time domain kernel p. Here, the secondary peak f(p) serves as a metric to measure the performance of PRT set. In next section, we will apply the DGPA to search the PRT set.

III. DISCRETE GENERALIZED PURSUIT ALGORITHMS

In this section, we will introduce the discrete generalized pursuit algorithm (DGPA). This algorithm is based on the learning automata [12] from the reinforcement learning. The aim of the DGPA is to find an optimal action out of the action space. The action space is composed of a set of actions F = [1, 2,…, F]. F is the size of the action space.

Besides, the DGPA has a probability vector defined by $P(t) = [p_1(t), p_2(t), \ldots, p_F(t)]$, where the $p_i(t)$ means the probability that the learning automata chooses the action i at the t-th iteration.

$$\sum_{i=1}^{F} p_i(t) = 1. \quad (11)$$

The probability vector P (t) is updated based on the reward estimation $d(t) = [d_1(t), d_2(t), \ldots, d_F(t)]$. The reward estimation d (t) is determined by the environment feedback [12].

In this paper, the definition of the action is a sequence consisting of ones and zeros, which is selected to find the optimal PRT set for PAPR reduction in OFDM systems. The length of the sequence S is N where M ones and N-M zeros are combined together. Here, the positions of the M ones in the sequence S are the PRT set which we need to find. The action is performed due to the minimum PAPR. Therefore, a certain action will get a positive reward from the environment feedback if this action is beneficial to the object function. On the contrary, the action receives a negative reward when this action has a bad influence on the object function.

The DGPA generalizes the method of the pursuit algorithm by pursuing all actions which have higher reward estimation than the current selected action. The actions which have higher reward estimation can acquire the increase of the probability. In the DGPA, the probability P(t) of the action is iteratively updated by the following equation:

$$P(t+1) = P(t) + \frac{\Delta}{K(t)} e(t) - \frac{\Delta}{F - K(t)} [u - e(t)], \quad (12)$$

where Δ is the resolution step.

$$\Delta = \frac{1}{F\delta}, \quad (13)$$

where the δ is the resolution parameter and F is the size of the action space.

At the t-th iteration, K(t) denotes the number of actions which have a higher reward estimation than the current selected action. u is a unique vector which is shown as

$$u_i = 1, i = 1, 2, ..., F. \quad (14)$$

e is a direction vector which follows the equations:

$$e_i(t) = \begin{cases} 1, if d_i(t) = \max\{d_i(t)\}, j \in \{1, 2.., F\} \\ 0, otherwise \end{cases}$$

$$e_j(t) = \begin{cases} 0, if d_j(t) \le d_i(t); \\ 1, if d_j(t) > d_i(t). \end{cases} \quad (15)$$

Based on (12) and (15), the probabilities of the selection action i and another action j are updated by the following equations.

$$\begin{cases} p_j(t+1) = \min\{p_j(t) + \frac{\Delta}{K(t)}, 1\}, if d_j(t) > d_i(t) \\ p_j(t+1) = \max\{p_j(t) - \frac{\Delta}{F - K(t)}, 0\}, if d_j(t) < d_i(t) \\ p_i(t+1) = 1 - \sum_{j \ne i} p_j(t+1) \end{cases} \quad (16)$$

Therefore, when the reward estimation of the action j is higher than that of the action i at the (t + 1)-th iteration, the probability of the action j will increase by

$$p_j(t+1) = p_j(t) + \frac{\Delta}{K(t)}, if d_j(t) > d_i(t) \quad (17)$$

where $p_j(t + 1)$ should be no more than 1.

On the contrary, when the reward estimation of the action j is less than that of the action i at the (t+1)-th iteration, the probability of the action j will decrease by

$$p_j(t+1) = p_j(t) - \frac{\Delta}{F - K(t)}, if d_j(t) < d_i(t), \quad (18)$$

where $p_j(t + 1)$ should be no less than 0.

To update the probability of each action, the reward estimation d(t) should be initialized firstly. Then for the selected action i, the reward estimation d(t) is updated by

$$\begin{cases} Z_i(t+1) = Z_i(t) + 1; \\ W_i(t+1) = W_i(t) + \beta(t); \\ d_i(t+1) = \frac{W_i(t+1)}{Z_i(t+1)} \end{cases} \quad (19)$$

where the $Z_i(t)$ means the number of times that the action I has been selected until the t-th iteration. The $W_i(t)$ denotes the number of times that the action i has been rewarded until the t-th iteration.

Here, β (t) represents the feedback from the environment. If the action i gets the positive feedback from the environment at the t-th iteration, β (t) is 1. Otherwise, the β(t) is 0 if the action i acquires the negative feedback from the environment.

That is to say, β (t) follows

$$\beta(t) = \begin{cases} 1, \\ 0, \end{cases} \text{the action } i \text{ gets the feedbacks.} \quad (20)$$

Therefore, if the feedback is positive, then the action i is rewarded. Oppositely, the action i is punished when the feedback is negative.

To conclude, the DGPA is presented in the **Algorithm 1**.

Algorithm 1 Discrete Generalized Pursuit Algorithm

Input: $F, \beta, \delta, Z_i, W_i, d_i, P_i, P_j$;
Output: the optimal action i;

Initialization: Randomly select the action i, set the probability of the action i as $P(0)$, record the feedback β, repeat until all actions are selected at least $Z_i(0)$ times.
Record the rewarded times of each action $W_i(0)$;
Calculate the $d_i(0)$, and $d_i(0) = W_i(0)/Z_i(0)$.
Learning Process
repeat
    Select the action i at the t-th iteration;
    Update the $d_i(t)$ by (19);
    Update the $P_i(t)$ by (16);
until the probability of action i is close to 1.

Repeat the ***Initialization*** *and* ***Learning Process*** *until the probability of action i is maximum*.

IV. DGPA-BASED PRT SET SEARCH

In this section, we apply the DGPA to find the PRT set.

In the TR based scheme, the action of finding the PRT set is defined to select the sequence of length N. In this sequence,

M is the length of PRT set and N-M is the length of zeros.

In our model, the TR scheme needs to generate the peak reduced signal c (k), k = 1, 2,…, N. Here, M numbers of zeros and N- M numbers of ones are composed of the peak reduced time signal c(k). The original time signal x(k) is combined with the peak reduced signal c(k) to generate the combined time signal a. Then the combined time signal a calculates its PAPR value. The objective of the optimization is to minimize the PAPR of the combined time signal a.

According to (9), the secondary peak f(p) is the evaluation of the performance for the peak reduced time signal. If the secondary peak f(p) increases after the (t+1)-th iteration, the action will get a negative feedback, which means that the β in (20) equals zero. However, the action can acquire a positive feedback when the secondary peak f(p) decreases after the (t + 1)-th iteration so that the β is one from (20).

In the initialization part, all actions are randomly selected firstly. The initialization step is ended until all actions are chosen at least Z(0) times. Then, the number of times that each action acquires the feedback is counted as $W_i(0)$, i = 1, 2,…, F, while the β is based on (20). And the reward estimation $d_i(0)$ is acquired by (19).

In the learning process, after getting β, the reward estimation d_i of the action i is calculated at each iteration based on (19). Then, the reward estimation d_i of each action is compared together to obtain the value of K.

For example, from the t-th iteration to the (t+1)-th iteration, the proposed method applies the DGPA to search the PRT set as follows.

1) Firstly, the value of β is taken as one or zero according to the change of secondary peak f(p).

2) Secondly, the action i has been selected $Z_i(t+1)$ times and rewarded $W_i(t + 1)$ times, respectively.

3) Thirdly, we calculate the reward estimation $d_i(t + 1)$ based on (19).

4) Fourthly, comparing the reward estimation of all actions, we obtain K(t+1) that the number of actions has higher reward estimation than the action i.

5) Fifthly, the probability of other actions, such as the probability $p_j(t+1)$ of the action j, is updated by (16).

6) Lastly, when the probability of any action approaches one, the algorithm ends and the number of this optimal action is output.

Here, the optimal action denotes the peak reduced time signal c which is related to the optimal PRT set. The optimal PRT set is the position indexes of ones in the peak reduced time signal c. In conclusion, the method of applying the DGPA to find the optimal PRT set is described above.

TABLE I. SIMULATION PARAMETERS

Parameters	Value
The number of sub-carriers N	512
The number of PRT set M	32
The number of OFDM symbols	1e5
The size of action space F	20
The resolution parameter δ	1
The number of selected times Z(0)	10

V. SIMULATION RESULTS AND ANALYSIS

In this section, simulation results of the proposed DGPA algorithm are presented in comparison with some baseline algorithms. We consider an OFDM system with N = 512 subcarriers and 16-QAM to be tested. The oversampling factor L is 4 in this paper. Unless otherwise specified, all the related parameters used in our simulations are summarized in the following Table I.

The PRT set obtained by the DGPA is the DGPA-PRT = {45, 85, 98, 103, 110, 114, 129, 138, 169, 213, 221, 233, 238, 243, 271, 300, 301, 314, 315, 335, 351, 362, 364, 368, 374, 391, 441, 445, 447, 460, 463, 489}.The secondary peak of the DGPA-PRT is 0.407.

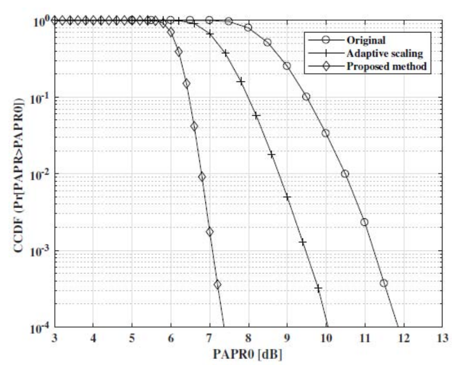

Figure 2. The comparison of CCDF between different methods with the same PRT.

In the Fig. 2, the PAPR reduction performance of proposed method is compared with AS-TR with the identical DGPAPRT. The AS-TR is referred to the paper [11]. Here, the maximum iteration is set 10 for AS-TR.

From the Fig. 2, when the CCDF equals 10^{-4}, the PAPR of original OFDM signal, AS-TR and the proposed method are 11.9 dB, 10.1 dB and 7.3 dB, respectively. Compared with the PAPR of original signal, the reduction gain is 4.5 dB obtained by the proposed method, which is 2.7 dB larger than AS-TR. It verifies that the proposed scheme with the DGPA-PRT can reduce PAPR better than AS-TR.

To evaluate the BER performance of the whole OFDM system, the processed signal passes through a solid-state power amplifier (SSPA) model,

$$S_o(n) = \frac{|S_i(n)|\, e^{j\theta n}}{[1+(\frac{|S_i(n)|}{C})^{2p}]^{\frac{1}{2p}}} \tag{21}$$

where $S_i(n)$ and $S_o(n)$ mean the input and output signals, respectively. P is 3 and C is 1.2. Then the processed signals are transmitted over the additive white Gaussian noise (AWGN) channel to calculate the BER.

Fig. 3 shows the comparison of BER performance between three different signals. For comparison, the BER of the original signal by an ideal SSPA with an infinite linear range is also plotted. This original linear signal has no distortion so that its BER performance is the best in three signals. From the Fig. 3, with the ideal SSPA, the original signal obtains the BER of 10^{-4} when Eb/No is 12.15 dB. However, with SSPA, for the proposed method and the AS-TR algorithm, Eb/No reaches 12.74 dB and 12.53 dB respectively when the BER is 10^{-4}.

Clearly, at the BER of 10^{-4}, the performance of proposed method is similar to those of original signal and AS-TR. Therefore, the proposed method can achieve the similar BER performance compared with AS-TR algorithm.

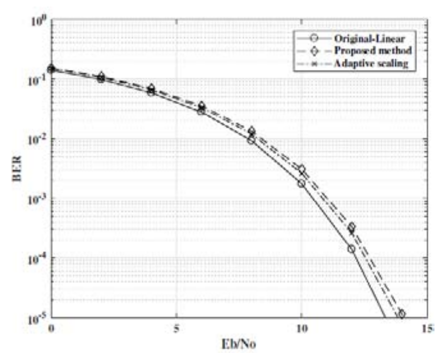

Figure 3. The comparison of BER performance between three signals.

VI. CONCLUSION

In this paper, we apply the DGPA to search the optimal PRT set based on TR method to reduce the PAPR in the OFDM systems. We first consider the minimum of the PAPR problem as our objective function and transfer this function to minimize the secondary peak in the TR method. Then we utilize the DGPA to search the PRT set for the PAPR reduction. The simulation results demonstrate that the proposed DGPA scheme can obtain a better PAPR reduction performance than other basic algorithms. Besides, the BER performance of the proposed method is similar to other benchmarks via comparison.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant 61371114 and 61501238.

REFERENCES

[1] T. Jiang and Y. Wu, "An Overview: Peak-to-Average Power Ratio Reduction Techniques for OFDM Signals," IEEE Trans. Broadcast., vol. 54, no. 2, pp. 257-268, June. 2008.

[2] X. Zhu, W. Pan, H. Li and Y. Tang, "Simplified Approach to Optimized Iterative Clipping and Filtering for PAPR Reduction of OFDM Signals," IEEE Transactions on Communications, vol. 61, no. 5, pp. 1891-1901, May 2013.

[3] D. Bi, P. Ren and Z. Xiang, "A Novel Joint PAPR Reduction Algorithm With Low Complexity Using LT Codes," IEEE Wireless Communications Letters, vol. 7, no. 2, pp. 166-169, April 2018.

[4] J. Y. Woo, H. S. Joo, K. H. Kim, J. S. No and D. J. Shin, "PAPR Analysis of Class-III SLM Scheme Based on Variance of Correlation of Alternative OFDM Signal Sequences," IEEE Communications Letters, vol. 19, no. 6, pp. 989-992, June 2015.

[5] H. S. Joo, K. H. Kim, J. S. No and D. J. Shin, "New PTS Schemes for PAPR Reduction of OFDM Signals Without Side Information," IEEE Transactions on Broadcasting, vol. 63, no. 3, pp. 562-570, Sept. 2017.

[6] J. Hou, C. Tellambura and J. Ge, "Tone injection for PAPR reduction using parallel tabu search algorithm in OFDM systems," IEEE Global Commun. Conf., Anaheim, CA, 2012, pp. 4899-4904.

[7] Y. Wang, W. Chen and C. Tellambura, "Genetic algorithm based nearly optimal peak reduction tone set selection for adaptive amplitude clipping PAPR reduction," IEEE Trans. Broadcast., vol. 58, pp. 462-471, Sep. 2012.

[8] S. H. Wang, W. L. Lin, B. R. Huang and C. P. Li, "PAPR Reduction in OFDM Systems Using Active Constellation Extension and Subcarrier Grouping Techniques," IEEE Communications Letters, vol. 20, no. 12, pp. 2378-2381, Dec. 2016.

[9] J. Tellado, "Peak to average power reduction for multicarrier Modulation," Ph.D. dissertation, Stanford Univ., 2000.

[10] J. C. Chen and C. P. Li, "Tone reservation using near-optimal peak reduction tone set selection algorithm for PAPR reduction in OFDM systems," IEEE Signal Processing Lett., vol. 17, pp.933-936, Nov, 2010.

[11] L. wang and C. Tellambura, "Analysis of clipping noise and tonereservation algorithms for peak reduction in OFDM systems," IEEE Trans. Veh Technol., vol. 57, no. 3, pp. 1675-1694, May. 2008.

[12] M. Agache, B. J. Oommen, "Generalized pursuit learning schemes: new families of continuous and discretized learning automata," IEEE Trans. Syst. Man Cybern. Part B Cybern., vol. 32, no. 6, pp. 738-749, Dec. 2002.

[13] C. Ma, M. Ding, H. Chen, Z. Lin, G. Mao, X. Li, "Socially aware distributed caching in device-to-device communication networks," Proc. 2016 IEEE Globecom Workshops, pp. 1-6, 2016.

[14] C. Ma et al., "Socially Aware Caching Strategy in Device-to-Device Communication Networks," IEEE Transactions on Vehicular Technology, vol. 67, no. 5, pp. 4615-4629, May 2018.

[15] L. Marini, J. Li, Y. Li, "Distributed caching based on decentralized learning automata," Proc. IEEE Int. Conf. Commun. IEEE, pp. 3807-3812, 2015.

[16] C. Ma, Z. Lin, L. Marini, J. Li and B. Vucetic, "Learning automaton based distributed caching for mobile social networks," Proc. IEEE Wireless Commun. Netw. Conf., pp. 1-6, 2016.