Ant-colony and nature-inspired heuristic models for NOMA systems: a review


1-8 Faculty of Engineering and Technology, Multimedia University, Malaysia
9 Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Malaysia

Article Info

Article history:
Received Dec 15, 2019
Revised Mar 17, 2020
Accepted Apr 7, 2020

Keywords:
Ant-colony
Heuristic
Orthogonal

ABSTRACT

The increasing computational complexity in scheduling the large number of users for non-orthogonal multiple access (NOMA) system and future cellular networks lead to the need for scheduling models with relatively lower computational complexity such as heuristic models. The main objective of this paper is to conduct a concise study on ant-colony optimization (ACO) methods and potential nature-inspired heuristic models for NOMA implementation in future high-speed networks. The issues, challenges and future work of ACO and other related heuristic models in NOMA are concisely reviewed. The throughput result of the proposed ACO method is observed to be close to the maximum theoretical value and stands 44% higher than that of the existing method. This result demonstrates the effectiveness of ACO implementation for NOMA user scheduling and grouping.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Hadhrami Ab. Ghani,
Faculty of Engineering and Technology,
Multimedia University,
Ayer Keroh Lama St., 75450 Bukit Beraung, Melaka, Malaysia.
Email: hadhrami.abdghani@mmu.edu.my

1. INTRODUCTION

Orthogonal multiple access (OMA) is a technique which allows multiple users to access wireless network bandwidth resources orthogonally such that the signal transmitted by the users will not interfere between each other over the frequency domain. On the other hand, non-orthogonal multiple access (NOMA) [1–6] is an alternative technique introduced to access wireless network bandwidth resources simultaneously whilst maintaining as well as improving the achievable capacity when power allocation is performed accordingly. Unlike OMA, NOMA promotes bandwidth sharing by multiple users to increase the bandwidth efficiency whilst increasing the sum capacity by sharing the same frequency carrier among users [7–11]. Sharing of bandwidth improves the spectral efficiency and more users can be accommodated at once. The capacity of the system can be enhanced by implementing NOMA by means of reducing the interference caused by the sharing user(s) in the same group via successive interference cancellation (SIC) [12–17]. Interference, without SIC, tend to grow worse when many users access the system at the same time. Therefore, user scheduling and allocation of resources along with SIC are necessary to reduce interference in the system which is caused by the sharing users.

The increasing number of users in cellular networks nowadays tends to level up the computational complexity [18] in determining the best user pairs or groups for accessing the system when these users are allocated with the same frequency bandwidth. While reducing the computational complexity, the achievable
throughput and spectral efficiency of NOMA should be maintained or improved to be better than OMA, which is the prior access technique implemented in cellular networks. To tackle these issues, scheduling methods with lower computational complexity including ACO-based schemes [19] and heuristic methods such as particle swarm optimization (PSO) schemes, genetic algorithms [20] are proposed for cellular and wireless networks [21] to improve the throughput, increase the spectral efficiency, and reduce the complexity and interference [22]. Other scheduling schemes have also been proposed in literature to obtain the best user pair such as round robin (RR). Scheduling but it has high computational complexity [18]. Hence, models with relatively lower computational load requirements such as heuristic artificial intelligence models are useful to be considered and studied in this paper in order to determine the user groups while improving the throughput and spectral efficiency as well as reducing the required computational complexity. Therefore the main contributions of this paper are firstly to carry out a concise study on the existing user scheduling schemes for NOMA based on ACO and other heuristic models and secondly to propose an ACO-based scheme which is demonstrated to be a suitable candidate for NOMA systems in achieving throughput and spectral efficiency with reduction in complexity.

2. LITERATURE REVIEW

NOMA is a technique which utilizes the radio access network (RAN) in order to provide radio connection between the mobile terminals and the radio network. In order to improve the capacity and reduce the latency, a suitable channel access method must be designed with the given available radio resources. As the demand for the capacity improvement with latency reduction keeps increasing, NOMA systems which is also offering higher bandwidth efficiency is a good candidate for achieving the target improvement. Various studies show that NOMA can be a good choice for channel access mechanisms. There are a number of desirable benefits from NOMA implementation including greater spectrum efficiency and sum capacity [23]. NOMA is generally implemented either as a power-domain system or code-domain system. In the power-domain system, the signals which come from multiple users are superposed to render a single resultant signal. This resultant signal is transmitted over the same channel. At the receiver side, the signals are de-multiplexed and detected with the help of multiuser detection (MUD) algorithms such as SIC [24]. Wireless transmission systems must accommodate an ultra-dense network with high number of users as it needs a reliable and fast channel access technique. One of the major challenges that will be faced by the system is the efficiency in terms of spectrum and energy. Due to the large number of signal transmissions occurring at the same time, they tend to experience different channel conditions and transmission requirements.

NOMA overcomes this problem by allocating the resources based on the quality of the user’s signal, which is typically measured by the signal-to-noise ratio. In other words, NOMA makes use its features to expand and contract the coverage of services based on the condition and demands from users all the time by operating at the right point which balances the spectrum and energy efficiency [25]. Although NOMA has greater spectrum efficiency, the system is exposed to interference from the sharing users [26]. Therefore, user scheduling is essential for reducing the interference and enhancing the bandwidth efficiency along with the successive interference cancellation, which will also contribute in eliminating the computational complexity of the algorithms [27]. Several methods are proposed for the formation of user pairs such as round robin [28]. As these approaches tend to require significantly higher computational loads when the number of users increases, new computationally lower schemes have been proposed in literature to determine the user pairs such as the heuristic models which are inspired by artificial intelligence approaches [29-34] which include drosophila optimization algorithm [35], particle swarm optimization algorithm [36-38], firefly optimization algorithm [39], dolphin echolocation algorithm [40], genetic algorithm [20, 41] and ant-colony optimization algorithm [19, 42]. These models have the ability to solve problems in varying fields such as, but not limited to, transportation, signal processing, image processing and biomedical engineering. This paper is presented focusing on the heuristic models which are inspired by natural phenomena.

3. NOMA SYSTEM MODEL

In this paper, the considered system model consists of a network having $N_s$ sites, each of which possesses a single cell that contains three sectors, where each of the sectors is represented as $A_{i,j}$ for $i \in [1, N_s]$ and $j \in [1, 3]$. Assuming that the first cell is located in the middle of the network, the analysis will be focused at this first cell without loss of generality. It is also assumed that all sectors in all sites of the networks are fully loaded, which means that the available frequency resources are fully utilized, uniformly and randomly, by all users in the sectors. Therefore, the focus of our analysis will be the first sector, denoted as $A_{1,1}$, of the first site located in the center of the network. There are $N_{us}$ users uniformly and randomly distributed in this sector. The users in each sector are allocated with $N_{rb}$ basic units of time-frequency resources known as physical resource blocks (RB) in the network.
For user $u$ in the $j$-th sector, $A_{i,j}$, of site $i$, the average received power over RB $r$ is expressed as follows [27]

$$P_{i,j,u,r} = P_u G_{p_i}(i,u) c_{i,u} G_A(i,j,u) f_{i,j,u,r},$$

where $P_u = \alpha_u P_T$ is the average transmit power, which is assumed to be equally allocated to each RB, for user $u$ over RB $r$ with $\alpha_u$ and $P_T$ are respectively the allocated power ratio and the total allocated power per RB. The following parameters, $G_{p_i}(i,u)$ and $c_{i,u}$, are the path gain and shadow fading between cell or site $i$ and user $u$, assuming that the shadowing effects experienced by the network sites are fully correlated. The antenna gain between sector $A_{i,j}$ and user $u$ is expressed as $G_A(i,j,u)$ and the small scale fast fading is represented as $f_{i,j,u,r}$ over RB $r$ between sector $A_{i,j}$ and user $u$. In a NOMA system generally implemented for user pairing, two users are configured to share the same RB $r$. In order to perform the SIC, the signal to interference plus noise ratio (SINR) of both users, user $u_1$ and user $u_2$ are measured for the case of no RB sharing which is the orthogonal multiple access (OMA) model, where the SINR $\gamma_{u_2,r}^{1,1}(u)$ for a user $u$ (which is applicable for both user $u_1$ and user $u_2$) at sector 1 of site 1 is written as follows

$$\gamma_{u_2,r}^{1,1}(u) = \frac{P_{1,1,u_2,r}}{\sum_{i=1}^{N_s} \sum_{j=1}^{N_g} P_{i,j,u_2,r} - P_{1,1,u_2,r}},$$

(1)

The noise expression can be excluded in the above equation as the considered system model is interference limited. If the SINR of user $u_1$ is larger than the SINR of user $u_2$ i.e. $\gamma_{u_1,r}^{1,1}(u_1) > \gamma_{u_2,r}^{1,1}(u_2)$ based on (1), then the average power allocation for both users $u_1$ and user $u_2$ in NOMA is $\alpha_1 P_T$ and $\alpha_2 P_T$ respectively as both users share the same RB $r$, where $\alpha_1 < 0.5$ and $\alpha_1 + \alpha_2 = 1$. Therefore, the SINR $\gamma_{u_2,r}^{1,1}(u_1, u_2)$ for user $u_2$ in sector 1 of site 1 can be expressed as follows

$$\gamma_{u_2,r}^{1,1}(u_1, u_2) = \frac{P_{1,1,u_2,r}}{\sum_{i=1}^{N_s} \sum_{j=1}^{N_g} P_{i,j,u_2,r} - P_{1,1,u_2,r} + P_{1,1,u_1,r}},$$

(2)

The SINR of user $u_2$ in (2) is determined directly without performing the SIC. In other words, decoding the received signal for user $u_2$ is performed directly without any prior SIC operation. Therefore, the interference caused by user $u_1$ must be included in the calculation of SINR of user $u_2$. In order to determine the SINR of user $u_1$, hence decoding the signal received for user $u_1$, the SIC operation is performed in order to remove the interference caused by user $u_2$. Therefore, no interference caused by user $u_2$ will be included in the expression of the SINR $\gamma_{u_1,r}^{1,1}(u_1, u_2)$ of user $u_1$, as written below

$$\gamma_{u_1,r}^{1,1}(u_1, u_2) = \frac{P_{1,1,u_1,r}}{\sum_{i=1}^{N_s} \sum_{j=1}^{N_g} P_{i,j,u_1,r} - P_{1,1,u_1,r}},$$

(3)

Therefore, the objective is to maximize the achievable throughput per group $R_{T,u,g}$ for each cell and site. For the case of cell $c$ and site $s$, the throughput per group $R_{T,u,g}$ to be maximized is

$$R_{T,u,g} = \sum_{n=1}^{N_g} \gamma_{u_1,r}^{c,s}(u_1, \ldots, u_{N_g}),$$

(4)

where the number of users per group is $N_g$. In the next section, the nature-inspired heuristic models for NOMA systems will be studied and compared against the proposed ACO scheme.

4. NATURE-INSPIRED HEURISTIC MODELS FOR NOMA SYSTEMS

Most of technologies which are created by human beings are inspired from nature such as insects’ behaviors [43]. This paper reviews this nature inspired heuristic models which are the candidates to be considered in finding a solution for getting the best user pairs in a wireless network.

4.1. Drosophila optimization model

Similar to ant-colony optimization, this method is inspired by the behaviour of drosophila of searching for food [35]. The concept of foraging behaviour is used for optimization. All drosophila fly to a particular location
once the location of fruit fly is found. Drosophila is able to detect the food location as it has a strong sense of smell and wide vision. It has the capability to look for food source within 25 miles. This optimization works by initializing the position of drosophila’s group based on range of variation for each individual. It follows by providing information on the direction and distance which are dependent on the characteristics of fruit fly when foraging. The flavour concentration determination value is then introduced.

The function of the density and the determination values are used to obtain the best flavour for drosophila’s group. At the same time, the optimum solution is discovered. Drosophila optimization is easy to be implemented when dealing with practical problems. It can be used to solve positive real problems with high precision. It has been demonstrated in [35] that Drosophila Optimization is useful for the large antenna systems, which are proposed to be one of the key features in 5G along with NOMA. However, when handling with non-positive (complex) problem, its stability seizes. Hence, users are not encouraged to apply this algorithm when dealing with negative arguments which cannot be interpreted or processed [36].

4.2. Particle swarm optimization model

Particle Swarm Optimization (PSO) is proposed in year 1995 by James Kennedy and Russell C. Eberhart. The principle of this method is wholly—based on the ability of groups. It is inspired by the concept of looking for food and process of courtship. The word ‘particle’ in PSO refers to bird [36]. The particle/bird switches its speed and direction based on information to maintain the optimum state and reach the right position. The solution is obtained only after a certain number of continuous iterations. For each cycle or iteration, two extremes are observed. The first extreme is known as ‘personal best’ which is identified by particles (bird) as optimal solution. The entire population observes the second extreme, the ‘global maximum’ as optimal solution.

These two parameters are important in determining the best solution/answer and speed in particular direction. PSO comprises of simple principles and less variables which make it is easy to be implemented. Moreover, it can be used in dealing with a wide range of issues such as non-linear, non-differentiable and so on [44]. In [45], PSO has been proposed to reduce the access delay in 5G systems. In addition to this, PSO has been demonstrated to improve power allocation in NOMA, as presented in [46]. Although it reduces the computational complexity required when non-heuristics models are used, no comparison has been made with other heuristic models such as ACO for NOMA systems. The number of iterations used in the proposed PSO model is relatively high, which goes up to 200 and the number of considered users are only between 2 to 4 users per sub-channel. Furthermore, PSO does not work for some problems like non-coordinate system.

4.3. Firefly optimization model

Likewise the heuristic algorithms mentioned above, this optimization tool is inspired by the behaviour of fireflies [36]. The flashing behaviour of fireflies is the underlying principle of this method. It consists of two parameters which evaluate and indicate the accuracy of the solution. The two indicators are fluorescence value and radius of perception. The fluorescence value is used to determine the quality of position for individual. On the other hand, the individual search is measured by the size of radius of perception. Initially, the individual (firefly) moves to the location of outstanding individual which is located within its area of search. The attractiveness of the brightness of firefly (due to fluorescence intensity) is essential to indicate the target. The location of target can be easily detected with better brightness. The attractiveness will be higher as the brightness is obvious. Thus, the target is clearly illustrated.

Factors such as fluorescence intensity and relative attractiveness must be addressed well when applying this technique to ensure the performance of the algorithm. The distance is inversely proportional to both the brightness and attractiveness. Firefly Optimization is easy to operate and it makes use of a small size of parameters. The negative point of this method takes a long time for converging and has a low probability of peak detection [36]. Firefly has also been employed in telecommunications, as studied in [39] for instance. Although firefly algorithm has been implemented for clustering sensor nodes or users in wireless sensor networks [47], the user grouping issue in NOMA has not been addressed.

4.4. Dolphin echolocation model

This approach is said to be similar to the rest of the optimization tools in as it shares some common characteristics. Therefore, the dolphin echolocation (DE) can be applied to find solutions for optimization issues. The concept of DE is based on the behaviour of a dolphin when foraging for prey. Initially, the dolphin looks for food within its search region and then focuses on a particular location once the target (prey) is found. It reduces the range of its search area as it approaches the prey. In other words, the echolocation process is proportional to the distance from the target. DE consists of two phases in its algorithm. Firstly, the dolphin initiates a global search where it searches the unexplored areas by choosing random paths within the search/look-for region and obtains some initial observation. In the second phase, the search range is limited or restricted to a smaller area and focuses on a particular location in a specific direction. This reduction of
look-for region is based on the results provided by the first phase. Most of the optimization tools apply these two phases in determining the optimal solution for a problem. The dolphin echolocation has the ability to adapt its algorithm according to the issues in which the technique is implemented. Hence, it can be used for solving various problems. In contrast, there is a high probability of information leakage occurrence in DE. Insufficient information affects the accuracy and stability of solution. This is due to the absence of essential data in determining and computing the optimal solution. Besides that, echolocation is only applicable within a limited range. Thus, it has to be implemented multiple times in different locations in order to have a wider range of search space [40, 48, 49]. Therefore, implementing DE in a user scheduling problem with a significantly wider search space such as NOMA with a large number of antennas is prohibitive.

4.5. Ant-colony optimization model

Ant-Colony optimization is an optimization tool which can be used for finding optimal or shortest paths based on the ants’ behaviour in searching for their food. An ant generally moves in a random fashion prior finding the food. When the food is found, the ant will return to its colony leaving markers, known as pheromones, that instigate other ants to take the same path for finding the food [19]. The concentration of pheromone in a particular (optimal) path depends on the number of ants using the same particular path to reach the food location. Thus, the evaporation rate of the pheromone will be low which in turn attracts more ants to use the path. In other words, ACO is a population-based meta-heuristic method that can be used to find an approximate solution to difficult optimization problems. The artificial ants look for a good solution to a given optimization problem. ACO is a strong technique which can be integrated with other heuristic algorithms to improve the performance of its own and other algorithms. Besides that, ACO has the ability to find a better solution in solving the problem. The proposed ACO algorithm [50] for NOMA is as follows

Input : The SINR of each user before sharing the bandwidth, the number of ants, $N_{ants}$, the number $N_u$ of users, the number $N_{ant}$ of users per group, the initial pheromone value, the pheromone value, the probability of choosing a user.

Output : The users in each group and the SINR of each user after sharing the bandwidth (NOMA)

Steps :

(i) The number of nodes of the ant colony optimization is set to $N_{ants}$ and the number of stages is also configured to $N_{ants}$.

(ii) Every ant will choose the next node, which is not yet chosen previously based on the initialized heuristic information, the pheromone value and its evaporation rate and the probability as described in [50].

(iii) The path with the maximum throughput per group, $R_{T,CS}$ as given in (4), which serves as the objective function, will be updated as well as the pheromone value, the evaporation rate and the probability before the next iteration is run.

(iv) The algorithm is terminated when the convergence criterion is fulfilled or the maximum number of iterations is reached.

Table 1 summarizes the pros and cons for all heuristics algorithms discussed in this paper. The best technique can be chosen based on the requirements set by the users. Since each of the heuristic models has plus points in different aspects, the best algorithm will be the one that suits the research scope and problems addressed. For grouping and scheduling the users in NOMA, the best heuristic algorithm presented in literature so far is ACO. It has been demonstrated to be able to group the users in cellular networks [19]. Other heuristic algorithms are suitable for other applications. PSO algorithms have been applied to improve power allocation in NOMA, as presented in [44-46]. Apart from improving the power allocation, PSO is also suitable for reducing the access delays in NOMA. However, the current general version of PSO is not suitable for grouping the users in NOMA.

<table>
<thead>
<tr>
<th>Nature-Inspired Algorithms</th>
<th>Advantage</th>
<th>Disadvantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ant-Colony Optimization [19]</td>
<td>Easy to be implemented</td>
<td>Requires longer search time</td>
</tr>
<tr>
<td></td>
<td>Ability to integrate with another</td>
<td></td>
</tr>
<tr>
<td></td>
<td>heuristic algorithm</td>
<td></td>
</tr>
<tr>
<td>Drosophila Optimization [35]</td>
<td>Deals with practical problems</td>
<td>Unable to use for non-positive issue</td>
</tr>
<tr>
<td></td>
<td>High precision</td>
<td></td>
</tr>
<tr>
<td>Particle Swarm Intelligence [36]</td>
<td>Easy to implement</td>
<td>Limitation on application of algorithm</td>
</tr>
<tr>
<td></td>
<td>Consists of less variables</td>
<td></td>
</tr>
<tr>
<td>Firefly Optimization [39]</td>
<td>Easy to operate</td>
<td>Low convergence rate</td>
</tr>
<tr>
<td></td>
<td>Less parameters involved</td>
<td>Low peak detection</td>
</tr>
<tr>
<td>Dolphin Echolocation [40]</td>
<td>Adaptation to various problems</td>
<td>Information leakage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Limited range</td>
</tr>
</tbody>
</table>
5. RESULTS

A NOMA system is considered to provide radio resources for users by grouping these users in pairs with one resource block allocated per pair or group. The number $N_{RB}$ of available resource blocks is varied from 1, 2, 3, 4 to 5 resource blocks with the number $N_U$ of available users varied as 2, 4, 6, 8 and 10 accordingly. In order to test the proposed ACO user grouping scheme, two other schemes are also executed which are the proportional fairness fixed allocation (PFFP) [27] and the exhaustive search (ES) method that serves as the upper bound since it finds the global optimum grouping. Figure 1 shows the mean throughput obtained when these approaches implemented in the NOMA system.

![Figure 1. The mean throughput in Mbps vs the number of resource blocks](image)

On the whole, the mean throughputs achieved by all three schemes under consideration increase as the number of resource blocks increases. When the number of resource blocks is between one and two, the mean throughput increases slowly and gradually. A more significant increase in the mean throughput can be observed when the number of resource blocks is between two and four. The gradient of the graph is the steepest when the number of resources falls in the range of four to five. It can be seen from this figure that the achievable mean throughput by ACO is higher than that of the PFFP. The gap between these two mean throughputs grows larger as the number of resource block increases. The mean throughput recorded for ACO is very close to the upper bound which is obtained via the exhaustive search (ES) method, where all possible combinations of user pairings are tested to find the maximum mean throughput. On the other hand, the PFFP method which is proposed in [27] to reduce the computational time and load, is observed to produce the lowest performance in terms of the achievable mean throughput. As a conclusion, the proposed ACO scheme has been demonstrated to provide the best user-pairs which maximizes the mean throughput, hence the spectral efficiency, close to the upper bound.

6. FUTURE WORK

This paper reviews some heuristic models which are potential to be regarded as candidate algorithms in NOMA for determining the best user groups. The main objectives of using this heuristic model are to reduce the computational load and increase the throughput of NOMA systems. Based on our concise review, ACO is a potential scheme to be considered for NOMA user scheduling as it has been demonstrated to be successfully employed for user grouping in cellular networks for the given radio bandwidth. Dolphin echolocation is not suitable as it is only applicable over a limited range of input space. However, ACO has no limitation on how many variables it has to work with. Firefly optimization and particle swarm optimization require a smaller number of parameters but so far implemented with a small number of users per sub channel. The practical solution provided by ACO [19, 42, 51] method for NOMA is useful as the number of mobile users using the cellular network services keeps increasing in this age. It is a method worth to be implemented and further developed and presented in the research and academic community. Therefore, a further study on ACO implementation with NOMA is the way forward to improve performance of user grouping along with other potentially effective improvements in the future.
7. CONCLUSION
Computational complexity in computing the best user pairs or groups becomes a barrier or challenge to achieve high throughput in NOMA. Therefore, ant-colony optimization (ACO) is considered to be applied in NOMA for solving the optimization problem. The increasing number of users, which will increase the required computational load without a good solution such as our proposed ACO-based approach, is an inevitable trend in this current age. As this ACO-based user grouping scheme has been proposed to perform well for the users sharing the same bandwidth in SC-FDMA and OFDMA networks, this scheme is potentially viable to be further developed for the future systems such as NOMA in 5G networks and beyond. Although the discussion and result in this paper is limited to two users per group for NOMA systems, further and future research can be carried out when the number of users is more than two per group. Therefore, employing ACO to user grouping for NOMA will certainly be useful for 5G networks and beyond.

ACKNOWLEDGEMENTS
The authors acknowledge and thank all the support from the Ministry of Education Malaysia for funding this research under the FRGS fund, with the grant code of FRGS/1/2019/TK04/MMU/02/2. Not to forget the support from Faculty of Engineering and Technology, Multimedia University, as well as all other individuals who are directly or indirectly involved in preparing this paper.

REFERENCES