

An Hybrid Approach for Job Scheduling in Multiclustering Environment Using BAT and CUCKOO Algorithm

Rajita Sharma

Dept. of Computer Science, Global Institutes, Amritsar, India-143001
 Email address: sharmarajita27[AT]gmail.com

Abstract— The Grid computing has increased huge consideration because of its innovative progression, cost lessening and accessibility. In the Grid situations, the appropriate scheduling of the got tasks over service gives has turned into a vital and essential issue. The scheduling issue in Grid situations is a NP-hard issue. The numerous heuristics have been proposed to take care of this issue up to now. In this paper, we propose another bee colony calculation to plan the tasks on service providers in the Grid conditions. The comes about exhibited that the proposed calculation has a superior operation as far as task execution time, holding up time and missed tasks.

Keywords— Job, BAT + CUCKOO algorithm, scheduling, federated grid, single processor, multi processor.

I. INTRODUCTION

In parallel computing multiple computing resources are used at same time to resolve a problem. In the parallel computing the large problem is broken into discrete parts that can be solved concurrently and each part is then further broken down into a sense of instruction. The instruction from each part is execute simultaneously on different processors and overall control/ coordination mechanism is employed.

“Parallel computing is a type of computing technique in which multiple processors execute an application or computation simultaneously. In the process of parallel computing large computations are splitted into smaller one and then workload between more than one processor is divided, the all work is computed at the same time”.

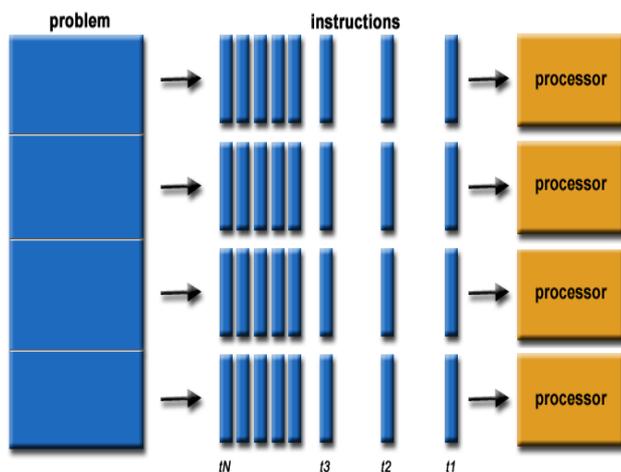


Fig. 1. Parallel computing architecture.

The enhancement of overhead as the load increases is accomplished through the proposed technique. The processor used within the parallel computation could be of similar types of dissimilar types. The similar type of processor used within the network form array processors and dissimilar types of processor within the network are known as vector processors. The processors can be used within the network executing instructions either in serial or parallel manner.

Grid computing is a novel framework and exceptionally famous wonder that spotlights on business asset arrangement and enables clients to use the computing assets introduced by various service providers(Sun et al. 2017). These foundations are spoken to by services that are utilized as well as introduced, conveyed or imitated with the assistance of virtualization. It is a model of service conveyance and get to system where virtualized assets are given as a service over the Internet. The primary objective of Grid computing is to give on-request computing services with high dependability, versatility, and accessibility in conveyed situations (Engineering 2013). It takes after a compensation for every utilization demonstrate and can be powerfully reconfigured to fulfill client demands through on-the-fly virtual assets. In the Grid computing, a few unique types of virtualized assets are scattered crosswise over numerous areas and systems. Four types of such virtualized assets are generally notable:

Programming as a Service (SaaS), Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Expert as a Service (EaaS). SaaS conveys uncommon reason programming that is remotely available by purchasers through the Web with an utilization based evaluating model(Nair 2016). IaaS gives equipment, programming, and gear to give programming application conditions, again with an utilization based estimating model. PaaS offers an abnormal state coordinated condition to assemble, test, and convey custom applications. EaaS gives the straightforward access to master, learning and abilities of HR remotely over the Internet. Task scheduler as a NP-Hard issue is an imperative piece of any appropriated framework like Grid, and P2P systems which allots employments to appropriate assets for execution. The objective of occupation scheduler is to limit the general execution time of a gathering of employments. The point of task scheduler in the Grid conditions is to abbreviate the reaction time and upgrade the service providers' usage. Due to the nature of scheduling issue (NP-Hard), numerous heuristics

have been proposed to take care of this issue up to now (Bouaziz and Chikhi 2014).

As another hand, simulated bee colony calculation is a swarm-based improvement strategy proposed for unraveling ceaseless improvement issues. It has been connected to take care of numerous issues and acquired intriguing comes about (Arvinder Kaur 2017). Be that as it may, to the best of our insight, no examination of its applications for task scheduling in the Grid situations has been done up to now. Along these lines, in this paper, we propose a task scheduler instrument to give the expressed points utilizing a counterfeit bee colony calculation.

Whatever is left of the paper is sorted out as takes after: In section 2, the related works and simulated bee colony calculation are quickly evaluated. In section 3, we propose a fake bee colony calculation to plan the employments aswell as the related calculations. In section 4 the acquired outcomes are introduced. In conclusion, conclusion and future works are given in section 5.

II. RELATED WORK AND BACKGROUND

This sections gives a concise outline of related looks into and foundation of the paper, including task scheduling in the Grid conditions and simulated bee colony calculations.

1.1 Task Scheduling in the Grid Environments

Job Scheduling

Scheduling refers to a set of policies which define the order of execution of processes. The scheduling in CPU is done to keep it as busy as possible. From all of the available resources of a computer system that needs scheduling before use, the CPU is one of the most critical of them. Multiprogramming is one of the basic and important scheduling technique. In parallel computing, multiple processors have to be scheduled, and it needs to manage the resources for all the processors.

1.2.1 Scheduling Policies in Parallel Computing

In parallel computing scheduling is difficult task as compare to serial computer. So there are some algorithm which is used for scheduling in parallel computing. Scheduling in parallel computing is NP-Complete.

The algorithm for parallel computing are:

- List Scheduling
- Look-ahead optimized scheduling(LOS)
- Longest processing time(LPT)

Some Heuristic Algorithm used for parallel computing:

- Opportunitunistic Load Balancing(OLB)
- User Directed Allocation(UDA)

1.2.2 Types of Multiprocessor Scheduling

a) *Thread Scheduling*: In thread scheduling priority assignments, thread dispatch, and thread migration are three main procedures. A periodic system interrupt is needed by scheduling algorithm to collect information of behavior of thread.

- *Priority Assignment*: There are two steps in priority in priority assignment procedures named as

synchronization sampling and priority assignment game

In Synchronization sampling a thread is responsible for a specific task, so there is a need to group the thread based on their types of task.

In Priority assignment game with the set of threads and their synchronization value priority to different threads is assigned in non-cooperative way. When a new thread is first enter the system the scheduling algorithm is assign an initial priority to it, then that scheduling algorithm perform synchronization sampling for it.

b) *Thread dispatch*: Entering of new thread in a system the operating system dispatch it to a processor with the shortest ready queue, that is the core with the smallest number of threads in its ready queue. In this way the priority of each thread could be different and the difference in the time slice between a high priority thread and low priority thread is very large.

c) *Thread migration*: When the ready queue of a core becomes free then that core is idle. The operating system scheduler uses a thread migration strategy to achieve better core utilization when it finds it finds the core has become idle.

1.2.3 Load Balancing

Load balancing algorithm can be categorized as static and dynamic. Static load balancing algorithms distribute the processes at compile time, while in the case of dynamic algorithm the processor processes at run time. Static load balancing algorithms rely on the estimate execution times of the processes and inter-process communication requirement. The parallel programs are dynamic and unpredictable in nature. In dynamic load balancing processes are generated and destroyed without a pattern at run time. Dynamic load balancing algorithm consist of four component that are Load measurement rule, an information Exchange rule, an initiation rule and a load balancing operation.

1.2.4 Approaches of Job Scheduling

1. *Centralized Approach*- It is an approach in which there is one main scheduler who schedules all the jobs present in a service queue to the resources present in a system. As one knows federated grid system consists of uncountable resources of varying environments, this approach is not suitable to manage all the resources and scheduling jobs to every other resource. It is highly prone to failure due to network congestion and moreover hardly reach to the idle resources in such a wide system network.

2. *Decentralized approach* enhances the utilization of resources by putting meta-scheduler at the top level of each grid system. Meta-scheduler maintains a queue on which number of jobs have to wait to be scheduled in a particular grid infrastructure. It is highly valuable in case of wide network to manage the resources by considering each specifications. It enhances the scalability of a system and if one system fails to performs its job, then it can easily transferred to the other system without much delay.

1.2.5 Approaches of Task Scheduling

These task scheduling approaches follow the decentralized workflow scheduling.

1. *Non-Coordinated Approach*- It is a approach in which the Meta-scheduler directly submits job to the resources without considering their load and utilization status.

2. *Coordinated Approach* is a approach in which the Meta-scheduler maintains the dynamic information of all the resources in a system before the allocation of job tasks.

The scheduling issue in the Grid conditions is more confused than that in conventional parallel frameworks, since Grid frameworks can have heterogeneity, dynamicity, irregular nearness, and huge correspondence overhead qualities. In this sub-segment, we quickly examination and abridge the best in class techniques and approaches in the field of task scheduling in the Grid situations.

(N. Kaur 2017) have proposed the close ideal scheduling arrangements to endeavor heterogeneity over various server farms for a Grid supplier. The strategy is viewed as various vitality productivity elements, (for example, vitality cost, carbon outflow rate, workload, and CPU control proficiency) which change crosswise over various information fixates relying upon their area, building outline, and administration framework. The have shown that the proposed technique can ready to accomplish all things considered up to 25% of vitality reserve funds in contrast with benefit based scheduling arrangements.

(Seleem, Mohamed, and Hussain 2013) have proposed another QoS-construct work process scheduling calculation situated in light of a novel idea called Partial Critical Paths (PCP), which tries to limit the cost of work process execution while meeting a client characterized due date. The proposed calculation recursively plans the PCP closes at already planned tasks. The outcomes exhibited that the calculation time of the proposed calculation is low for the taken a toll diminishing and the reasonable strategies, yet is any longer for the advanced arrangement.

(N. Kaur 2017) A scheduling calculation to address the significant difficulties of task scheduling in the Grid situations is proposed by Choudhary and Peddoju. In this calculation, the approaching tasks are gathered on the premise of the task necessity like least execution time or least cost and organized. At that point, asset choice is done on the premise of task requirements utilizing a ravenous approach.

(Aggarwal, Rawat, and Upadhyay) A four-level engineering for multilingual data assets scheduling in the Grid situations is proposed by Han and Luo in 2013. It incorporates client getting to level, innovation supporting level, asset scheduling level and assets level. They proposed a three-layer scheduling model for multilingual data assets in the Grid computing. The model incorporates some home chiefs, some neighborhood scheduling specialists and a worldwide scheduling operator. The illustration demonstrated that the proposed scheduling model could enhance the execution of the scheduling of multilingual data assets in the Grid conditions.

(Sejpal 2016) have streamlined the occupation scheduling utilizing biogeography-based improvement (BBO). BBO relocation is utilized to change existing arrangements and to adjust new great arrangements. BBO offers the benefit of versatile process, which is produced for double number occupation scheduling issue in Grid computing. Test comes

about demonstrated that the execution of the proposed techniques are superior to the considered other strategies in the occupation scheduling issues.

(Amandeep Kaur 2017) have proposed a model which considers information administration to get tasteful makespan on various server farms. In the meantime, their versatile information reliance investigation can uncover parallelization openings. They presents a versatile technique for work process applications. It comprise of a set-up stage which assembles the groups for the work process tasks and datasets, and a run-time arrange which makes the covered execution for the work processes. Through thorough execution assessment examines, they showed that the proposed technique can adequately enhance the work process finish time and usage of assets in a Grid condition.

(Anon 2016) At long last, JafariNavimipour and SharifiMilani have proposed another calculation in view of cuckoo pursuit calculation to plan the tasks in Grid computing which depends on the commit brood parasitic conduct of some cuckoo species in blend with the Lévy flight conduct of a few feathered creatures and organic product flies. The reenactment comes about showed that when the estimation of Pa is low, the speed and scope of the calculation turn out to be high.

Artificial Bee Colony Algorithm

The manufactured bee colony calculation as a novel meta-heuristic approach which was created in 2005(Dash and Mohanty 2014). It is roused by the smart scavenging conduct of bumble bees and is the reenactment of the moderate scrounging model of bumble bee in scan handle for settling genuine parameter, non-arched, and non-smooth advancement issues. In this calculation, a colony comprises of three sorts of simulated bees including utilized, spectator furthermore, Cuckoos.

Employed bee: Each sustenance source is related with an utilized bee and each utilized bee tries to identify a new sustenance source in the area of its present nourishment source. The recognized sustenance source is retained when the nectar measure of the distinguished nourishment source is higher than the nectar measure of current sustenance source. After culmination of the inquiry procedure, utilized bees share their data concerning the nectar measure of nourishment sources with passerby bees through waggle move in the move zone (Alam and Optimization 2016).

BAT bee: They watches the move of utilized bees and picks one of their sources relying upon the moves, and after that goes to that source. In the wake of picking a neighbor around that, she assesses its nectar sum (Cheung, Ding, and Shen 2016). An passer by bee assesses the data picked up from the utilized bees and tries to locate another sustenance source in the neighborhood of the chose sustenance in light of this assessed data. Along these lines, the propensity of passerby bees is to look around the nourishment sources with high nectar sum; along these lines, more qualified sustenance sources can be picked for misuse.

Cuckoo: The quantity of Cuckoos is not predefined in the colony. A Cuckoo is created by the circumstance of a nourishment source whether it is surrendered or not. Cuckoos

are deciphered from a couple utilized bees, which surrender their sustenance sources through a foreordained number of cycles (Li, Ma, and Yang 2017). The quantity of Cuckoos is controlled by "as far as possible" speaking to the quantity of trials before characterizing a nourishment source as "surrendered". In the event that a sustenance source can't be enhanced amid the foreordained number of unsuccessful trials being equivalent to utmost parameter, that nourishment source is leaved and another sustenance source is haphazardly created. Investigation what's more, misuse methodology are done together for hearty inquiry handle in this module [36].

In the manufactured bee colony demonstrate, a sustenance source signifies a conceivable arrangement of an advancement issue and the nectar measure of a nourishment source speaks to the nature of an answer. As specified some time recently, every sustenance source is related with an utilized bee in the colony. In this manner, the quantity of nourishment sources is equivalent to the quantity of utilized bees. Additionally, the quantity of utilized bees is equivalent to the quantity of passerby bees. In this manner, the number of nourishment sources is half of the population number.

III. PROPOSED METHOD

In the proposed technique, the position of a nourishment source speaks to a conceivable answer for the scheduling issue furthermore, the nectar measure of a nourishment source compares to the quality (wellness) of the related answer for given scheduling issue. The quantity of the utilized bees is equivalent to the quantity of arrangements in the populace. At the initial step, a haphazardly appropriated introductory populace is produced. After instatement, the populace is subjected to rehash the cycles of the inquiry procedures of the utilized, passerby, and Cuckoos, separately. An utilizedbee creates an alteration on the source position in her memory and finds another sustenance source position. Given that the nectar measure of the new one is higher than that of the past source, the bee remembers the new source position and overlooks the old one. Else she keeps the position of the one in her memory. All things considered utilized bees finish the inquiry procedure, they share the position data of the sources with the spectators on the move range. Every passerby assesses the nectar data taken from every single utilized bee and after that picks a nourishment source contingent upon the nectar measures of sources. As on account of the utilized bee, she delivers a change on the source position in her memory and checks its nectar sum. Giving that its nectar is higher than that of the past one, the bee remembers the new position and overlooks the old one. The sources relinquished are resolved and new sources are haphazardly created to be supplanted with the deserted ones by fake scouts [38].

After introduction, the populace is subjected to rehash the cycles of the pursuit procedures of the utilized, passerby, and Cuckoos, individually. An utilized bee creates an adjustment on the source position in her memory and finds another nourishment source position. Given that the nectar measure of the new one is higher than that of the past source, the bee remembers the new source position and overlooks the old one. Else she keeps the position of the one in her memory. After

every utilized bee finish the pursuit procedure, they share the position data of the sources with the spectators on the move territory. Every spectator assesses the nectar data taken from every single utilized bee and afterward picks a sustenance source contingent upon the nectar measures of sources. As on account of the utilized bee, she delivers a change on the source position in her memory furthermore, checks its nectar sum. Giving that its nectar is higher than that of the past one, the bee remembers the new position and overlooks the old one. The sources relinquished are resolved and new sources are arbitrarily created to be supplanted with the surrendered ones by manufactured scouts. The means of the calculation are:

- Produce starting populace.
- While stop criteria is not fulfilled, perform steps 3 to steps 6.
- Send the utilized bees onto their nourishment sources.
- Send the passer by bees onto the nourishment sources relying upon their nectar sums.
- Send the Cuckoos to pursuit conceivable new sustenance sources.
- Memorize the best nourishment source discovered up until this point.

We consider t tasks to be prepared on c Grid hubs. There are a few suppositions and obliges as takes after: each task has predefined number of operations and a known decided arrangement among these operations; each Grid hub is prepared at zero time; each task must be handled on one machine at a given time; and each Grid hub can handle another operation simply in the wake of finishing the antecedent operation. Give C_{ti} a chance to be the finishing time of task T_i . At that point, the fulfillment time (makespan) is:

$$F = \max \{C_{ti} \mid i = 1, \dots, n\}$$

3.1 Initialization Phase

In this stage, new competitor arrangements are created for each utilized bees. Toward the starting, the underlying BsCuckoos are put haphazardly in Grid hubs and B_s is the quantity of Cuckoos. With the assistance of the randomization and the addition based nearby inquiry, the underlying populace is developed with both differing qualities and quality.

3.2. Employed Bee Phase

To improve the procedure of calculation, the quantity of utilized bees is set to the same as that of sustenance sources, also, each utilized bee is connected to every arrangement in the populace, individually. Additionally, an utilized bee produces a sustenance source in the area of the present source.

3.3. BAT Phase

We right off the bat give a joined nearby pursuit which consolidates the inclusion based neighborhood look as well as a swap-based neighborhood look. Every passerby bee picks an utilized bee with a specific end goal to enhance its answer. This determination is finished by wellness estimations of utilized bees by roulette wheel

$$(\sum f_j n_j = 1).$$

3.4. Cuckoo Phase

The surrender counters of every single utilized bee are tried. The utilized bee, which can't enhance self-solution until the surrender counter reaches as far as possible, progresses toward becoming Cuckoo. The Cuckoo turns into the utilized bee. In this manner, Cuckoos avert latency of utilized bee populace. Along these lines, Cuckoos with the most noteworthy wellness are picked as "Chosen Bee" and the went by locales by them are browsed neighbourhood of Grid hubs.

IV. EXPERIMENTAL RESULTS

Keeping in mind the end goal to test the execution of the proposed strategy, we utilize a standard assessment procedure in the Grid conditions. In this area, we examine whether our approach is viable or not. We display the subtle elements of broad examinations to assess the execution of the approach. The recreation is performed utilizing CPU center i5, memory 4GByte and 4MByte store memory. The outcomes have been acquired with a framework test system modified in MatlabR2013b1. As another hand, creating irregular DAGs, enables us to assess diverse application charts. We had assessed the execution of the proposed calculation under various parameters, including diverse quantities of subtasks. The proposed calculation is contrasted and different calculations as far as makespan. The quantity of created subtasks in a DAG is chosen from 1 to 20. The computational what's more, correspondence expenses of the DAG are created arbitrarily from a range.

TABLE I. Makespan of schedule.

Process	Existing	Proposed
10	1440	500
20	1450	520
30	1460	540
40	1470	570
50	1480	600

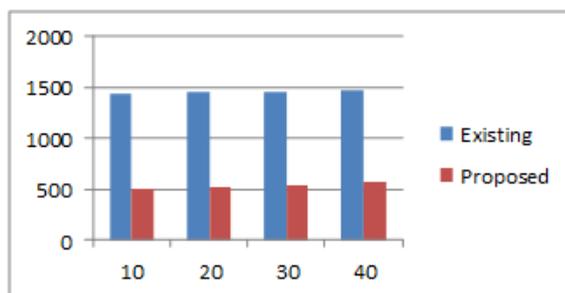


Fig. 2. Plot of makespan of schedule.

The flowtime associated with the given schedule is as under

TABLE II. Showing flow time of schedule.

Process	Existing	Proposed
10	500	105
20	530	135
30	545	145
40	560	160
50	580	175
60	600	200

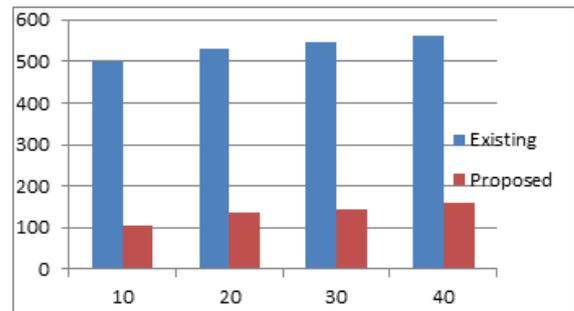


Fig. 3. Showing flowtime of schedule.

X - Exis define the no, of jobs

Y- Exis dfine the Makespanm and Flowtime

V. CONCLUSION AND FUTURE WORKS

Grid computing is a novel framework and exceptionally prevalent marvel that spotlights on business asset arrangement and enables clients to used the computing assets facilitated by various service providers. Task scheduling is one of the essential issue in the Grid conditions which is gotten numerous considerations these days in light of its high calculation cost. Consequently, numerous heuristics have been proposed to take care of this issue up to now. Due to the favorable of a fake bee colony calculation, in this paper, we propose another counterfeit bee colony calculation to plan the tasks on service providers in the Grid situations.

REFERENCES

- [1] A. Aggarwal, T. K. Rawat, and D. K. Upadhyay, "Optimal design of L 1 -Norm based IIR digital differentiators and integrators using the bat algorithm," vol. 5, pp. 1-14.
- [2] Alam, Shah and A. Single Objective Optimization. 2016. "Assessment of evolutionary programming, firefly algorithm and cuckoo search algorithm in single- objective optimization," (December):16-18.
- [3] Anon. 2016. "Sensing in Sparse Satellite Cognitive Systems."
- [4] A. Bouaziz, A. Draa, and S. Chikhi, "A cuckoo search algorithm for fingerprint image contrast enhancement," *Second World Conference on Complex Systems (WCCS)*, 2014.
- [5] N. J. Cheung, X.-M. Ding, and H.-B. Shen, "A nonhomogeneous cuckoo search algorithm based on quantum mechanism for real parameter optimization," *IEEE Transactions on Cybernetics*, pp. 1-12, 2016.
- [6] M. Dash, and R. Mohanty, "Cuckoo search algorithm for speech recognition," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 3, issue 10, pp. 3540-3545, 2014.
- [7] M. RamaMohana Rao and A. V. Naresh Babu, "Optimal power flow using cuckoo optimization algorithm," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 2, issue 9, pp. 4213-4218, 2013.
- [8] A. Kaur, "A review on various job scheduling algorithms," vol. 13, issue 3, pp. 359-67, 2017.
- [9] A. Kaur and A. P. Agrawal, "A comparative study of bat and cuckoo search algorithm for regression test case selection," *7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, 2017.
- [10] N. Kaur and A. Chhabra, "Comparative analysis of job scheduling algorithms in parallel and distributed computing environments," *International Journal of Advanced Research in Computer Science*, vol. 8, issue 3, pp. 948-956, 2017.
- [11] X. Li, Xiangtao, S. Ma, and G. Yang, "Synthesis of difference patterns for monopulse antennas by an improved cuckoo search algorithm," *IEEE Antennas and Wireless Propagation Letters*, vol. 16, issue 3, pp. 141-44, 2017.
- [12] S. S. Nair, K. P. S. Rana, and V. Kumar, "Efficient 2nd order full band digital IIR integrator design using cuckoo search algorithm," *2nd*



International Conference on Communication Control and Intelligent Systems (CCIS), 2016.

- [13] S. Sejpal and N. Shah, "A novel multiple objective optimized color watermarking scheme based on LWT-SVD domain using nature based bat algorithm and firefly algorithm," *IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT)*, pp. 38-44, 2016.
- [14] Y. Z. Seleem, M. H. Mohamed, and K. F. Hussain, "Improving genetic process mining using honey bee algorithm," *2nd International Conference on Informatics and Applications (ICIA)*, pp. 59-65, 2013.
- [15] G. Sun, Y. Liu, Z. Chen, Y. Zhang, A. Wang, and S. Liang, "Thinning of concentric circular antenna arrays using improved discrete cuckoo search algorithm," *IEEE Wireless Communications and Networking Conference (WCNC)*, 2017.