International Journal of Computer Applications (0975 – 8887)
Volume 73– No.10, July 2013

Ant Colony Optimization: A Swarm Intelligence based Technique

Manju
Research Scholar
Dept. of Computer Sci. & Application,
Kurukshetra University, Kurukshetra

Chander Kant
Assistant Professor
Dept. of Computer Sci. & Application,
Kurukshetra University, Kurukshetra

ABSTRACT

The growing complexity of real-world problems has motivated computer scientists to search for efficient problem-solving methods. Divide and conquer techniques are one way to solve large and difficult problems. Division of large work into smaller parts and combining the solution of small problems to get the solution of large one has been a practice in computer research since long time. Swarm also exhibits the behavior of division of work and cooperation to achieve difficult tasks. Evolutionary computation and swarm intelligence meta-heuristics are outstanding examples which show that nature has been an unending source of inspiration. Artificial Swarm/Ant foraging utilizes various forms of indirect communication, involving the implicit transfer of information from agent to agent through modification of the environment. Using this approach, one can design efficient searching methods that can find solution to complex optimization problems. Over times, several algorithms have been designed and used that are inspired by the foraging behavior of real ants colonies to find solutions to difficult problems. In this paper the idea of Ant Colonies is presented with brief introduction to its applications in different areas of problem solving in computer science.

General Terms
Artificial Intelligence, Swarm Intelligence, Data Mining.

Keywords
Swarm Intelligence, Ant Colony Optimization, Association Rule Mining.

1. INTRODUCTION

Multi-Agent research draws inspiration from the behavior of insects as they provide a good example of cooperation and self- organization. Insects have relatively simpler behavior individually, but with the amazing capability of organizing their actions, they represent a complex and highly structured social organization. Swarm intelligence is a revolutionary technique for solving optimization problems that formerly took its inspiration from the biological examples that can be observed in nature, such as ant colonies, flocks of birds, fish schools and bee hives, where a number of individuals with limited capabilities are able to come to intelligent solutions for complex problems [1]. Swarm intelligence is a relatively new subfield of artificial intelligence which studies the emergent collective intelligence of groups of simple agents. Swarms inherently use forms of decentralized control and self-organization to achieve their goals. Taking inspiration from the success and efficiency of swarms in real world, computer researchers have tried to develop sophisticated methods and systems that make use of the techniques of the swarms to find solutions to difficult problems. In biological swarms, the individuals (ant, bee, termite, bird or fish) are not complete engineers, but instead are simple creatures with limited cognitive abilities and limited means to communicate. Yet the complete swarm exhibits intelligent or collective behavior, providing efficient solutions for complex problems such as predator evasion and shortest path finding. The main principles of the collective behavior as presented in Fig 1 are:

a) *Homogeneity*: every agent in flock has the same behavior model. The agent moves without a leader, even though temporary leaders seem to appear.
b) *Locality*: the motion of each agent is only influenced by its nearest flock mates. Vision is considered to be the most important senses for flock organization.

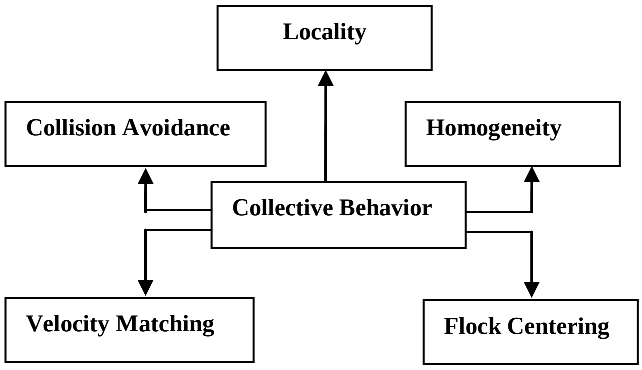

Fig 1: Main Principles of Collective Behavior of Swarms

c) *Collision Avoidance*: avoid with nearby flock mates.
d) *Velocity Matching:* attempt to match velocity with nearby flock mates.
e) *Flock Centering*: attempt to stay close to nearby flock mates.

2. RELATED WORK

In the beginning, the two mainstreams of the swarm intelligence area were: Ant Colony Optimization (Dorigo and Stützle, 2004) [1] and Particle Swarm Optimization (Kennedy and Eberhart, 1995) [7].

The ant colony optimization (ACO) meta-heuristics is inspired by the foraging behavior of ants. The ants' goal is to find the shortest path between a food source and the nest. Each path constructed by the ants represents a potential solution to the problem being solved. When foraging for food, ants lay down a chemical substance called pheromone. Ants can locally communicate to each other by means of the pheromone trails deposited in the environment. As more ants use a path, more pheromone will be deposited, and consequently, more ants will be attracted to it and leading them to find the shortest path between a food source and the colony.

The PSO meta-heuristics is motivated by the coordinate movement of fish schools and bird flocks. The PSO is compounded by a swarm of particles. Each particle represents a potential solution to the problem being solved and the position of a particle is determined by the solution it currently represents.

In recent years, new swarm intelligence algorithms have appeared that are:

In Bee foraging, honey bee colonies have a decentralized system to collect the food and can adjust the searching pattern precisely in order to enhance the collection of nectar. Bees can estimate the distance from the hive to food sources by measuring the amount of energy consumed when they fly, besides the direction and the quality of the food source. This information is shared with their nest mates by performing a waggle dance and direct contact [6].

Bacterial foraging (BFO), applied the algorithm to the optimization of a benchmark function. In reproduction, the health of each bacterium represents its fitness value. All bacteria are sorted according to their health status and only the first half of population survives. The surviving bacteria are split into two identical ones in order to form a new population. Thus, the population of bacteria is kept constant [8].

In Lampyridae bioluminescence, Lampyridae is a family of insects that are capable to produce natural light (bioluminescence) to attract a mate or a prey. They are commonly called fireflies or lightning bugs. The firefly algorithm (FA) was proposed by Yang [9] and the algorithm was applied to the optimization of benchmark functions.

In RIO algorithm, cockroaches' agents are defined using three simple behaviors: cockroaches search for the darkest location in the search space and the fitness value is directly proportional to the level of darkness (find darkness phase); cockroaches socialize with nearby cockroaches (find friend phase) and third behavior is cockroaches periodically become hungry and leave the friendship to search for food (find food phase) [5].

In Bat Algorithm (BA), based on the echolocation behaviour of bats. It is potentially more powerful than particle swarm optimization and genetic algorithms. The primary reason is that BA uses a good combination of major advantages of these algorithms in some way. In these studies, highly be needed to carry out the sensitivity analysis, to analyze the rate of algorithm convergence, and to improve the convergence rate even further [10].

The shuffled frog-leaping algorithm is a memetic metaheuristic that is designed to seek a global optimal solution by performing a heuristic search. It is based on the evolution of memes carried by individuals and a global exchange of information among the population. it combines the benefits of the local search tool of the particle swarm optimization and the idea of mixing information from parallel local searches to move toward a global solution [13].

3. ANT COLONY SYSTEM

Ant Colony Optimization (ACO) was introduced in the early 1990s as a novel technique to solve hard combinatorial optimization problems. ACO is an optimization algorithm inspired in the collective foraging behavior of ants to find and exploit the food source that is nearest to the nest. ACO is based on cooperative search paradigm that is applicable to the solution of combinatorial optimization problem. Communications among individuals or between individuals and the surroundings is based on the use of chemicals formed by the ants called pheromone [2]. The different types of Ant Colony models are Ant system for Traveling Salesman Problem (TSP), Ant Miner, Ant quality, Ant-Cycle, Max-Min ACS, Ant System (AS), ranking model and Ant density . The design of an ACO algorithm implies the specification of the following aspects:

a) A proper definition of the problem that needs to be solved by the ants so that some strategies can be made to incrementally build solution to the problem based on some transition rule, amount of pheromone on the path and some local information.
b) A method to construct valid solutions that can be considered legally permissible in the real world situation.
c) A heuristic function that should be designed based on the problem to measure the significance of the item/terms that can be added to the current solution so as to give right direction to the process of finding optimal solution.
d) A rule for updating the pheromones on certain paths that leads to good solution to the problem and fine-tune the pheromone-trail.
e) A transition rule that is based on the heuristic function and amount of pheromone on a path to decide about the movement of ants and construction of solution.

3.1 Real Ants System

The visual sensitivity of many ant species is not as much developed and there are ants that are completely blind. The communication between them is based on the use of chemicals produced by them. These chemicals are called Pheromones. While walking from the food source to the nest or vice-versa, ants deposit pheromones on the ground, forming a pheromone trail. Ants can smell the pheromone and they tend to choose, probabilistically, paths marked by strong pheromone concentrations. So more the ants follow a trail, the more attractive that trail becomes to be followed by other ants[3]. The basic idea was designed and run by Deneubourg et. al. in 1990 [12], who used a double bridge connecting a food source and the nest. One of the bridges was twice as longer as the other. Initially all the ants choose the outgoing path randomly. Some ants choose the short path and reach the food earlier than the other ants choosing the longer one. The ants reaching the destination through the short path returned to the nest before the others on the long path. Thus the amount of pheromone on the smaller path tend to increase faster than the longer path and results an increase in the probability of choosing that path by the foragers. Since ants prefer to follow trails with larger amounts of pheromones, eventually all the ants converge to the shorter path.

3.2 Artificial Ants

An artificial ant simulates a real ant and a set of artificial ants develops mechanisms of cooperation and learning. Ant Colony Optimization (ACO) was developed by Dorigo et. al. in 1996 [1] and it was first used to solve combinatorial optimization problems. Artificial ants are characterized as agents that imitate the behavior of real ants. However it should be noted that an artificial ant system has some differences in comparison with real ants which are as follows:

a) Artificial ants have memory.
b) They are not completely blind.
c) They follow a discrete time system.

The main idea is that when an ant has to select a path among several available paths, the ant chooses the one which is chosen more frequently by other ants in the past. Thus path with larger amount of pheromones is the shorter path and chosen by most of the ants. The Ant System works in two major steps:

a) Construction of the solution to the problem under consideration.
b) Updating the pheromone trails which may increase or decrease the amount of pheromone on certain paths.

3.2.1 *Solution Construction*

A solution is constructed in a probabilistic way. Initially a fixed amount of pheromone s deposited on each path. At the very first stage each ant starts randomly. The selection of next path to be followed or an item to be added to the current solution is determined by the ants based on two major things:

a) A problem dependent heuristic function that may determine the quality or predictive power of the item under consideration.
b) The amount of pheromones deposited by ants in the past.

3.2.2 *Pheromone Updation*

Initially equal amount of pheromone is deposited on each path. As each ant moves, it tries to build a best possible solution to the problem. When the ant completes construction of the solution, the amount of pheromone must be updated on each item in the path. It can be done as follows:

3.2.2.1 Increasing the amount on selected path

When an ant completes its trip, the quality of the solution constructed is checked. The increase in pheromone can be proportional to the quality of the solution; better the solution, larger the increase in the amount of pheromone. This corresponds to increasing the chances of selecting the items in the current solution by other ants in the future.

3.2.2.2 Decreasing the amount of pheromone

In the real ant system, the pheromone tend to evaporate with time, thus the artificial ant system also needs to decrease the amount of pheromone on the paths that are not part of the constructed solution. As per the need, the decreasing rate can be set to a constant or it can be determined by the frequency of the items used during past.

4. APPLICATION AREAS FOR ACO AND CURRENT TRENDS

ACO is true meta-heuristic with dozens of application areas. Inspite a massive boost in the performance of ACO algorithms and the theoretical understanding about the working of the system, there are several areas where only primarily steps have been taken and there is a vast scope for research. The major application areas of Ant Colony Optimization include the following [3]:

a) Routing- This area include the problems in which one or more agents have to visit a predefined set of locations and whose objective function depends on the ordering in which locations are visited. This area includes Travelling Salesman Problem, sequential Ordering problem, vehicle routing etc.

b) Assignment- The task in the assignment problems is to assign a set of items to a given number of resources subject to some constraints. This may include Quadratic Assignment Problem, Generalized Assignment Problem, Frequency Assignment Problem, Graph Coloring Problem, and University Course Timetabling Problem.

c) Scheduling- Scheduling in broad sense is concerned with allocation of scarce resources to tasks over time. Scheduling problems mainly occurs in production and manufacturing industries and include Single- Machine total tardiness Scheduling Problem, Job Shop, Open Shop and Group Shop Scheduling Problem, Resource-Constrained Project Scheduling, Permutation Flow Shop Problem etc.

d) Subset- In subset problems, a solution to the problem under consideration is represented as a subset of the set of available items subject to problem- specific constraints. This include set Covering Problem, Graph Tree partition Problem, Arc Weighted l-Cardinality Tree Problem, Multiple Knapsack Problem, Maximum independent Set problem, Redundancy Allocation Problem etc.

e) Machine Learning- This is a new area of optimization problems where ACO algorithms can be applied to solve extremely hard problems. This includes generation of classical Rules, data Clustering, Regression and several others.ACO has been applied to a few such problems so far, opening the application of this algorithmic technique to a new range of important problems.

f) Network routing- This area includes the problems of finding minimum cost path among all pairs of the nodes in the network. Here if the cost of links associated with the nodes varies dynamically with network topology and type of communication networks, thus rendering the problem of routing too complex to be solved by simple techniques such as Dijkstra's algorithm. ACO algorithms opened new ways to tackle such problems.

5. SCOPE OF ACO IN RULE MINING

Data mining technology aim to find useful patterns from large amount of data. Data mining is the process of analyzing data from different angles & getting useful information about data. The data mining can help in predicting a trend or values, classifying, categorizing the data & in finding correlations, patterns from the dataset. The overall goal of data mining process is to extract information from a dataset & transform it into an understandable structure for future use. Association rule mining is a one of the most important technique in data mining. It extracts significant patterns from transaction databases and generates rules used in many decision support application [15].

5.1 Association Rules in Data Mining

Consider I = $\{i_{1,} i_2 \dots i_n\}$ be a set of items. Let D be a set of transactions or database. Each transaction t∊D is an item set such that t is a proper subset of I. A transaction t supports X, a set of items in I, if X is a proper subset of t. An association rule is an implication of the form X→Y, where X and Y are subsets of I and X∩Y= Ø. The support of rule X→Y can be computed by the following equation: Support (X→Y) = |X→Y| / |D|, where |X→Y| denotes the number of transactions in the database that contains the itemset XY, and |D| denotes the number of the transactions in the database D. The confidence of rule is calculated by following equation: Confidence (X→Y) = |X→Y|/|X|, where |X| is number of transactions in database D that contains item set X. A rule XY is strong if support(X→Y) ≥ min_support and confidence(X→Y) ≥ min_confidence, where min_support and min_confidence are two given minimum thresholds. Association rule mining algorithms scan the database of transactions and calculate the support and confidence of the rules and retrieve only those rules having support and confidence higher than the user specified minimum support and confidence threshold [11], Association rule mining consists of two stages viz. the discovery of frequent itemsets and the generation of association rules. It follows, that in the vast majority of cases, the discovery of the frequent set dominates the performance of the whole process. Therefore, we explicitly focus the paper on the discovery of such set [14].

There are a large number of approaches for finding interesting association rules such as AIS (Agrawal, Imielinski, Swami) algorithm [11], Apriori algorithm, Dynamic Itemset counting (DIC) algorithm, Pincer-Search algorithm, Frequent Pattern (FP) Tree and many more. Most of these algorithms rely on scanning the database time and again to find frequent patterns. Thus large number of passes over the entire database consumes a lot of time and the performance of the algorithm degrades.

5.2 ACO & Rule Mining

As ACO is a true meta-heuristic that can search for a solution to optimization problem very quickly, it can be applied to search interesting association rules, classification rules and other patterns from the large databases. Here the problem is to find frequent patterns that are interesting and have sufficient amount of confidence. An efficient algorithm can be designed based on Ant colony system in which a number of ants traverse the database and generate rules at the end of each pass. The system parameters such as the choice of pheromone and the heuristic function can be set such that the generated rules are of good quality. Other models of ant system such as brood sorting, division of labor can be used to boost the performance of the algorithm.

6. CONCLUSION

This paper presents a brief outline of the concept of Ant Colony Systems and swarm intelligence with a review of its projected application areas. The scope of ACO in the field of data mining for finding new and useful patterns quickly is also highlighted. The technique of ACO can be applied to optimization problems to achieve a high quality solution. The area of data mining is a new application of ACO and still a lot of work has to be done in this field. Data mining tasks such as clustering, association rule mining, regression analysis etc. can make extensive use of this approach to find quality solutions.

7. REFERENCES

[1] Dorigo, M., and Stützle, T. "Ant Colony Optimization", MIT Press, 2004Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[2] Parpinelli, R.S., and Lopes, H.S. 2011 "New inspirations in swarm intelligence: a survey", International Journal of Bio-Inspired Computation, Vol. 3, No. 1, pp. 1-16.

[3] Dorigo, M., Maniezzo, V. & Colorni, A. 1996 " Ant System: Optimization by a Colony of Cooperating Agents", IEEE Transactions on Systems, Man, and Cybernetics-Part B, Vol. 2, No. 1, pp. 29-41.

[4] Fraga, H. 2008 "Firefly luminescence: a historical perspective and recent developments", Journal of Photochemical & Photobiological Sciences, Vol. 7, pp. 146–158.

[5] Havens, T., Spain, C., Salmon, N. and Keller, J. 2008 "Roach infestation optimization", IEEE Swarm Intelligence Symposium, September, pp.1–7.

[6] Karaboga, D. 2005 "An idea based on honey bee swarm for numerical optimization', Technical report, Erciyes University, Engineering Faculty, Computer Engineering Department.

[7] Kennedy, J., and Eberhart, R.C. 1995 "Particle swarm optimization", Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ., pp. 1942-1948.

[8] Passino, K. 2002 "Biomimicry of bacterial foraging for distributed optimization and control", IEEE Control Systems Magazine, pp. 52–67.

[9] Yang, X.S. 2009 "Firefly algorithms for multimodal optimization", SAGA 2009, Lecture Notes in Computer Science, 5792, pp. 169-178.

[10] Yang, X. 2010 "A new metaheuristic bat-inspired algorithm", Nature Inspired Cooerative Strategies for Optimization (NISCO 2010), Studies in Computational Intelligence, Vol. 284, pp. 65–74, Springer Berlin.

[11] Agrawal, R., Imielinski, T., and Swami, A. 1993 "Mining association rules between sets of items in large databases", Proceedings of the ACM SIGMOD Conference on Management of Data, pp. 207-216.

[12] Deneubourg, J. -L., Aron, S., Goss, S., & Pasteels, J.M. 1990 "The Self –Organizing Exploratory Pattern of the Argentine Ant:, Journal of Insect Behaviour, Vol. 3, pp. 159-168.

[13] Eusuff, M.M. and Lansey, K.E. 2003 "Optimization of water distribution network design using the shuffled frog leaping algorithm", Journal of Water Resource Planning Management, , 210 – 225.Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), pp. 1289-1305.

[14] Dao-I Lin, Zvi M Kedem. 1997 "Pincer Search: A New Algorithm for Discovering the Maximum Frequent Item Set", New York University.J. Han, M. Kamber. 2001 "Data Mining: Concepts and Techniques" Morgan Kaufmann Publishers.