Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/090119

A Modified Artificial Bee Colony Algorithm for Solving Least-Cost Path Problem in Raster GIS

Khalid A. Eldrandaly, Mohammed M. Hassan and Nabil M. AbdelAziz*

Information Systems Department, Faculty of Computers and Informatics, Zagazig University, Egypt

Received: 15 Nov.2013, Revised: 6 Feb. 2014, Accepted: 10 Feb. 2014
Published online: 1 Jan. 2015

Abstract: The computation of least-cost paths over a cost surface is a well-known and widely used capability of raster geographic information systems (GISs). It consists of finding the path with the lowest accumulated cost between two locations in a raster model of a cost surface. This paper presents a modified Artificial Bee Colony (ABC) algorithm for solving least-cost path problem in a raster-based GIS. This modification includes the incorporation of a distinct feature which is not present in the classical ABC. A new component, the direction guidance search method, is used to guide a bee walking toward the final destination more efficiently. In addition, this paper examines how the quality of the raster-based paths can be improved by using larger connectivity patterns. The experimental results show that the performance of the modified ABC model is quite close to Dijkstra's algorithm while its computational complexity and solution time is much lower than Dijkstra's algorithm. The results also, indicate that raster-based paths can be improved by using larger connectivity patterns.

Keywords: Least-Cost Path, Cost Surface, Raster Data Modeling, Swarm Intelligence, Artificial Bee Colony Algorithm.

1 Introduction

Finding a least-cost-path in a raster data format is a very important function in geographic information systems [1]. Least cost path analysis consists of finding the path with the lowest accumulated cost between two locations in a raster model of a cost surface [2]. It has been used to solve many real-world problems [3], such as finding the best traversal paths across off-road terrain [4,5], or the alignment of linear constructions, like roads [1], canals [6], trails [7], power lines [8], and pipelines [9]. This method has also been applied in ecology [10], archeology [11], and public health [12], to name just a few examples. The cost surface is represented by a raster map in which each cell is associated with a cost score, reflecting the impedance of movement per a distance unit across the cell [3,8]. Dijkstra's algorithm is one of the most commonly used algorithms for solving the least cost path problem [1,3,8]. It was designed for tracing the shortest path in a network with nodes connected by weighted links. To use this algorithm in a raster-based GIS, a virtual network can be constructed in which the centers of each raster cell serve as the nodes in the network, and the connections between the neighboring cell centers act as the links of the network [1]. In most GIS software, connectivity is established within the 3 * 3 cell neighborhood matrix, resulting in eight movement angles (four in cardinal and another four in diagonal directions) from a focal cell [3]. Unfortunately, paths calculated using Dijkstra's (or related) algorithm are subject to two main drawbacks [3,13]: one is the distortion inherently present in raster-based paths, and the other is the high computational effort needed to calculate the paths, especially when large rasters are concerned. The use of larger neighborhoods is one conventional strategy to decrease the distortion in paths induced by the raster structure [3]. Instead of the conventional eight-connected pattern, larger patterns embracing 16, 32 or even more neighbors have been used [3,14,15]. Figure 1 shows the allowed directions of movement from a focal cell associated with different raster-based neighborhood patterns. The redundant neighbors indicated in white are excluded. Even though distortion can be decreased in this manner, it comes with the cost of increasingly dense graphs and associated computational expense [13,16].

Recently, natural swarm intelligence has been used to tackle a variety of complex computation problems, such as functional optimization, route finding, scheduling, structural optimization, vehicle routing and image and

* Corresponding author e-mail: khalid_eldrandaly@zu.edu.eg

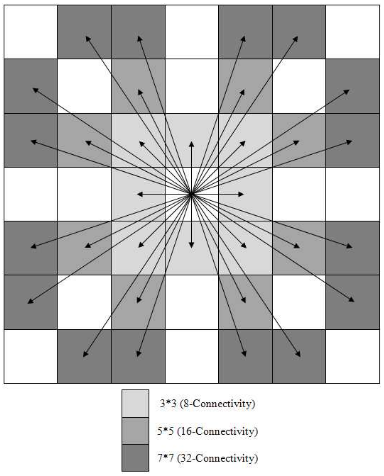

Fig. 1: Neighborhood patterns (adapted from [3])

data analysis [17]. The swarm intelligence is usually designed by imitating the flocking of birds or the swarming of insects such as bees and ants [18]. Artificial Bee Colony (ABC), developed by Karaboga in 2005 [19], is one of the most widely used swarm intelligence techniques for solving real world problems [20,21,22]. ABC algorithm is a nature-inspired metaheuristic, which imitates the foraging behavior of bees. It has been tailored successfully, to solve a wide variety of discrete and continuous optimization problems, and revealed promising results in processing time and solution quality [20,22]. According to the best of our knowledge ABC has not been used yet for solving least-cost path problem in a raster-based GIS.

This paper presents a modified Artificial Bee Colony (ABC) algorithm for solving least-cost path problem in a raster-based GIS. This modification includes the incorporation of a distinct feature which is not present in the classical ABC. A new component, the direction guidance search method, is used to guide a bee walking toward the final destination more efficiently. In addition, this paper examines how the quality of the raster-based paths can be improved by using larger connectivity patterns.

2 ABC Algorithm for solving least-cost path problem

2.1 The Conventional ABC algorithm

In 2005, Karaboga analyzes the foraging behavior of honey bee swarm and proposes a new algorithm simulating this behavior for solving multi-dimensional and multi-modal optimization problems, called Artificial Bee Colony [23]. The artificial bee colony consists of three groups of bees [21]: employed bees associated with specific food sources, onlooker bees watching the dance of employed bees within the hive to choose a food source, and scout bees searching for food sources randomly. Both onlookers and scouts are also called unemployed bees. Initially, all food source positions are discovered by scout bees. Thereafter, the nectar of food sources are exploited by employed bees and onlooker bees, and this continual exploitation will ultimately cause them to become exhausted. Then, the employed bee whose food source has been exhausted becomes a scout bee, and their food sources are abandoned. In ABC, the position of a food source represents a possible solution to the problem, the nectar amount of a food source corresponds to the quality (fitness) of the associated solution and the number of employed bees is equal to the number of food sources (solutions). The search carried out by the artificial bees can be summarized as follows [24]: (1) employed bees determine a food source within the neighborhood of the food source in their memory, (2) employed bees share their information with onlookers within the hive and then the onlookers select one of the food sources, (3) onlookers select a food source within the neighborhood of the food sources chosen by themselves, and (4) an employed bee of which the source has been abandoned becomes a scout and starts to search a new food source randomly.

2.2 Modifying ABC for adapting to least-cost path problem

2.2.1 Proposed Direction guidance search method

ABC algorithm seems to be straightforward for solving least cost path problem in a raster-based GIS because of bees' exploration capability. In the optimization process, an artificial bee can visit any cell on raster surfaces during path exploration. A bee can move randomly on the two-dimensional raster surface if there are no constraints and regulations. During bees' moving, there are 8, 16 and 32 direct neighboring cells from a central cell on a raster surface based on 3x3, 5x5 and 7x7 neighborhood patterns respectively (Figure 1). These neighbors represent possible moving directions. There are infinite combinations of the moving schemes by forming a path between an origin and a destination. At first, the cells of the path are selected randomly by starting from the start cell to the end cell. New cells should be selected from the neighboring cells in order not to break off the route (break off control). The generation process stops when the destination cell is reached. In every selection, the control of re-selection is made in order not to use the same cell again. Using this random selection method, infinite loops have emerged and sometimes the route started to rotate around itself like a snake (Figure 2). Thus, a bee could move randomly without completing a path because of lacking the vision capability. Therefore, conventional

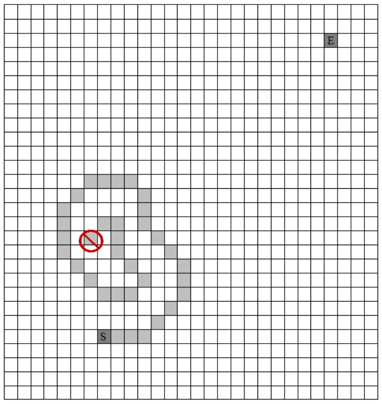

Fig. 2: Path exploration using the conventional ABC algorithm (3*3 neighborhood pattern 8-connectivity)

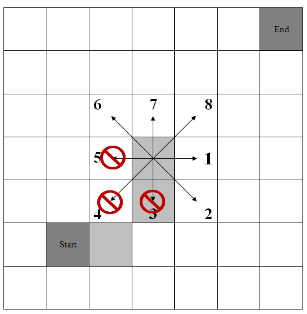

Fig. 3: Example of restricted directions (3*3 neighborhood pattern 8-connectivity)

ABC algorithm should be modified for adapting to the least cost path problem by incorporating a more sophisticated direction guidance search method to provide some 'vision' capability.

In the proposed direction guidance search method the bee is encouraged to move towards its final destination during path exploration via determining some restricted directions that distract the bee from the destination cell. The restricted directions are determined using the angular calculation between the current cell and the destination cell. All cells located in the opposite quarter of the calculated angle are specified as restricted directions (Figure 3). Then in every new cell generation, the control of re-selection is made besides the break off control. With the help of the direction guidance search method, unnecessary generation of the cells is prevented and the speed of the generation is accelerated. Pseudo-code of the direction guidance search method is shown in figure 4

```
Direction Guidance Search (Neighborhood Pattern, Start cell, End cell) {
(1) Add start cell to path
(2) Calculate the angle between start cell and end cell
(3) According to the angle calculated and the neighborhood pattern
    determine the restricted direction
(4) According to neighborhood pattern generate a random number
    represents the movement direction.
        If (3*3 neighborhood pattern) random number between 1-8;
        If (5*5 neighborhood pattern) random number between 1-16;
        If (7*7 neighborhood pattern) random number between 1-32;
    remove restricted direction.
(5) According to movement direction the new cell is defined
(6) If (New cell <> End cell)
        Direction Guidance Search (Neighborhood Pattern, New cell, End cell)
(7) Return path }
```

Fig. 4: Pseudo-code of the direction guidance search method.

2.2.2 Modified ABC Algorithm

In the modified ABC algorithm, the path (food source) consists of connected cells according to predefined neighborhood pattern, that is, a sequence of connected cells from the starting cell (source) to the end cell (destination). Each cell has (x, y) coordinates and the cost of travelling through that cell. The accumulative cost of the path (food source) is the nectar amount. The objective function is to minimize the accumulative cost of the path (nectar amount).

In the initialization phase and after determining the neighborhood pattern, the population of food sources (different paths) is initialized by artificial scout bees and control parameters are set. The scouts start to explore the environment with the help of the direction guidance search method in order to find a food source. For each bee, the start cell and the end cell are the same, but the intermediate cells are diverse. Then the population of food sources (different paths) is evaluated by calculating by the nectar amount (accumulative cost).

In the employed bees phase, artificial employed bees search for new path within the neighborhood of the current path in their memory. They find a neighbor path by determining a random part of path and replace it with new part using direction guidance search method and then evaluate its fitness (accumulative cost). After producing the new path, its fitness is calculated and a greedy selection is applied between it and its parent, where the best of the two paths are selected with probability P_{better} and the worst of the two with probability ($1 - P_{better}$). After that, employed bees share their path information with onlooker bees waiting in the hive by dancing on the dancing area.

In the onlooker bees phase, artificial onlooker bees probabilistically choose their path depending on the information provided by the employed bees. For this purpose, a stochastic selection scheme based on the fitness (nectar) values, which is similar to roulette wheel selection, is used. After a path for an onlooker bee is probabilistically chosen, a neighborhood path is determined using direction guidance search method, and its fitness value is computed. As in the employed bees phase, a greedy selection is applied between two sources.

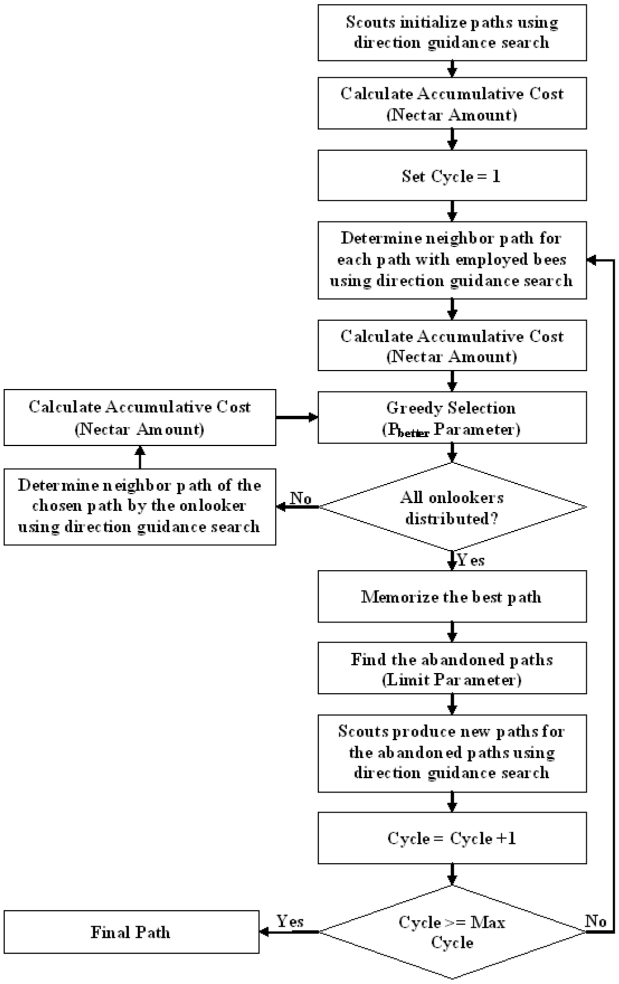

Fig. 5: Flowchart of the Proposed ABC Algorithm for Least Cost Path Problem

In the scout bees phase, employed bees whose paths cannot be improved through a predetermined number of trials, called limit, become scouts and their paths are abandoned. Then, the scouts start to search for new paths using direction guidance search method. Hence, those paths which are initially poor or have been made poor by exploitation are abandoned.

Employed bees phase, Onlooker bees phase, and Scout bees phase are repeated until a termination criteria is satisfied such as maximum cycle number. Figure 5 represent the flowchart of proposed ABC algorithm for solving least cost path problem in a raster-based GIS.

3 Model implementation and Validation

In order to evaluate the performance of the modified ABC algorithm, it has been tested on solving a hypothetical case study using a real raster dataset and the results obtained are compared with that of the state-of-the art algorithm (Dijkstra's algorithm). In this case study, the town of Stowe, Vermont, USA [25] would like to find the best route for a new access road from the new school location site to a nearby road intersection. A cost raster dataset (767x767 cells) is used for solving this problem (Figure 6).

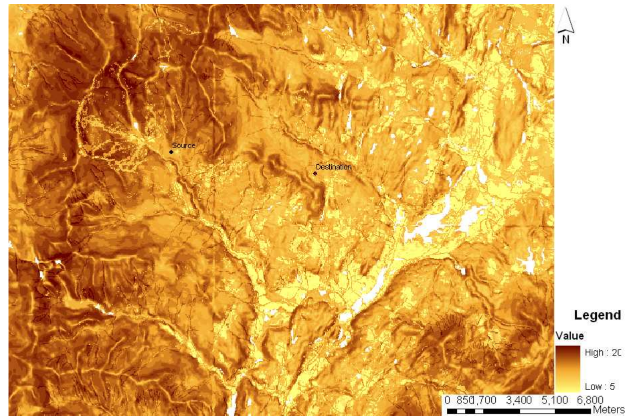

Fig. 6: The raster dataset used in the case study

Least cost paths are calculated using both the conventional Dijkstra's algorithm and the proposed ABC algorithm with three different neighborhood patterns (3 x 3, 5 x 5, and 7 x 7). The two algorithms were evaluated in two aspects: the total cost distance of the paths, and the time needed to calculate them. The cost distance can be considered to imply the quality of the paths. A relatively short cost distance of a path calculated by a certain algorithm indicates that the algorithm in question is capable of determining paths without unnecessarily strong distortions in them, or that it is capable of finding better routes for the paths altogether. The second aspect of the evaluation is the processing time required to calculate paths. The processing time considered here only includes the time required to determine of the least-cost paths from a source node until the destination node is found. The two algorithms (Dijkstra and modified ABC) were implemented in ArcGIS software using C# programming language and the calculations were done using a laptop computer with 2.00 GB of RAM and a 2.00 GHz Intel Core 2 Duo Processor.

3.1 Setting the ABC parameters

One of the major steps in preparing to use the proposed ABC algorithm is the setting of ABC parameters such as total number of bees, Maximum number of cycles, limit and P_{better}. Several experiments with 3 x 3, 5 x 5, and 7 x 7 cell neighborhoods were carried out to determine the proper values of these parameters for solving the current case study. The results of these experiments are shown in Table 1 and Figure 7.

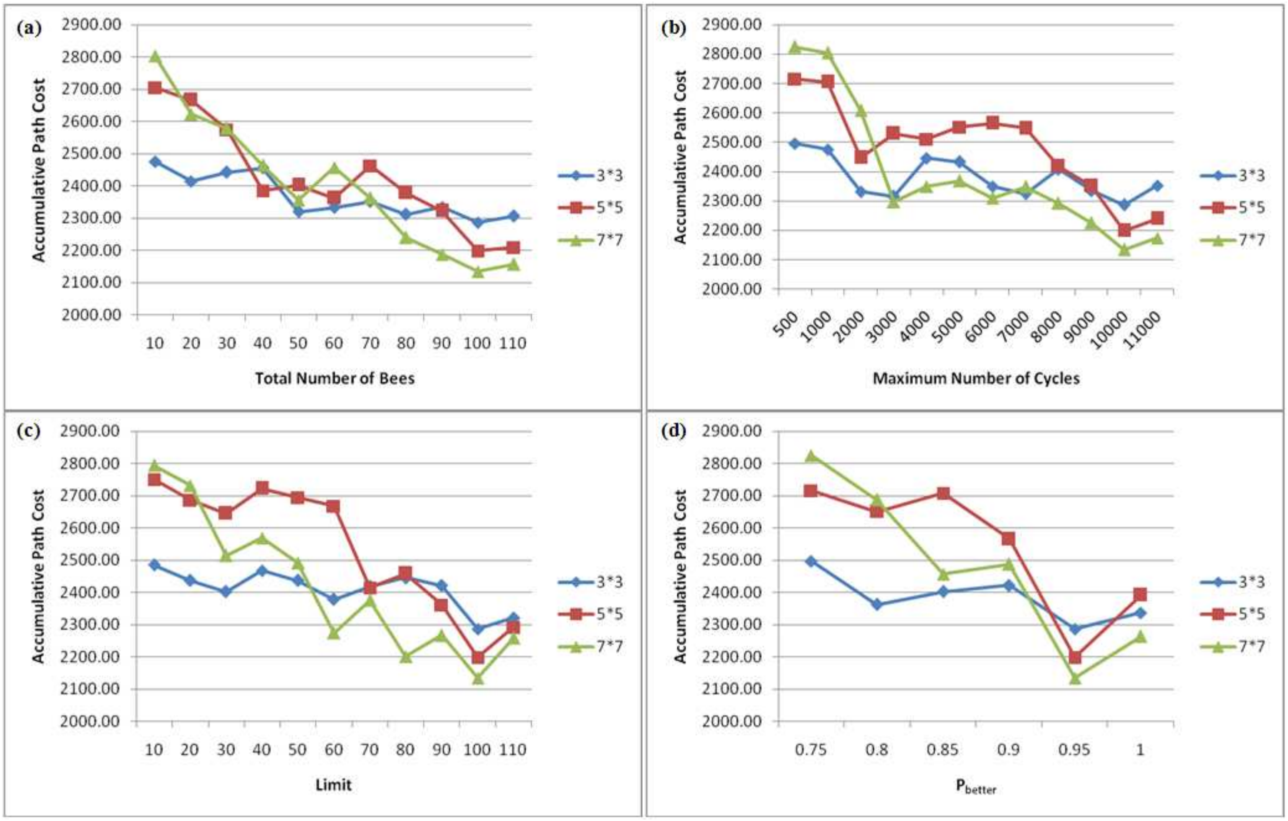

Fig. 7: a) Accumulative Path Cost vs. Total Number of Bees, b) Accumulative Path Cost vs. Maximum Number of Cycles, c) Accumulative Path Cost vs. Limit, and d) Accumulative Path Cost vs. P_{better}

Table 1: Parameters to be used in the proposed ABC model

Parameters	Value
Total number of bees	100
Maximum number of cycles	10000
Limit	100
P_{better}	95%

3.2 Stability Validation of the proposed algorithm

Heuristic algorithms are subject to some uncertainties thus the stability of repeated simulations is an important indicator for assessing the validity of the proposed model [18]. Figure 8 is the overlay of the optimal paths from 10 repeated simulations. It clearly shows that the proposed ABC algorithm can repeat the simulation results although there are some minor differences. Moreover, the simulations can be also repeated by changing the moving direction. There are two different types of moving forward (from the source to the destination) and backward (from the destination to the source). It is expected that the simulated patterns should be very similar between the forward moving and backward moving. This assumption has been confirmed by the experiment (figure 9). Table 2 also compares the accumulative costs obtained from 10 different simulations using both the forward and backward moving with different neighbor patterns. The variations are very small because these two types of walking yield the similar mean values and small standard deviation values.

3.3 Evaluation of the proposed model

To evaluate the performance of the proposed ABC algorithm, its results are compared with that of the Dijkstra's algorithm as shown in Table 3 and Figure 10. While Table 3 summarizes the total cost distance of the paths, and the time needed to calculate them using both the Dijkstra's algorithm and the modified ABC algorithm, Figure 10 shows the two least-cost paths identified by using the two algorithms.

Figure 10 and Table 3 clearly show that the proposed ABC can find the near optimal solution with error no more than 2.5%, while it takes 20% of the Dijkstra's algorithm time. The results confirm the expectations of the use of large search neighborhoods, that is, a large number of connections, producing better paths. Although

Table 2: The accumulative costs obtained from 10 different simulations.

Neighborhood Pattern	3*3		5*5		7*7	
Method	Forward moving	Backward moving	Forward moving	Backward moving	Forward moving	Backward moving
Sim1	2321.42	2351.43	2368.35	2339.92	2166.64	2156.55
Sim2	2280.54	2384.44	2268.12	2228.45	2228.45	2145.00
Sim3	2342.04	2358.96	2237.20	2275.79	2119.48	2160.61
Sim4	2295.70	2366.11	2215.60	2320.58	2100.14	2116.41
Sim5	2445.65	2346.29	2301.28	2325.93	2124.07	2254.07
Sim6	2381.34	2395.63	2279.30	2267.56	2089.48	2141.32
Sim7	2253.96	2415.21	2308.46	2373.30	2188.20	2159.38
Sim8	2409.87	2375.32	2295.12	2213.84	2168.64	2212.62
Sim9	2322.65	2328.84	2218.18	2224.13	2149.44	2090.50
Sim10	2327.06	2263.06	2260.35	2351.44	2132.32	2146.17
Mean	2338.02	2358.53	2275.19	2292.09	2146.69	2158.26
Std	59.10	41.89	46.42	57.76	42.42	46.06

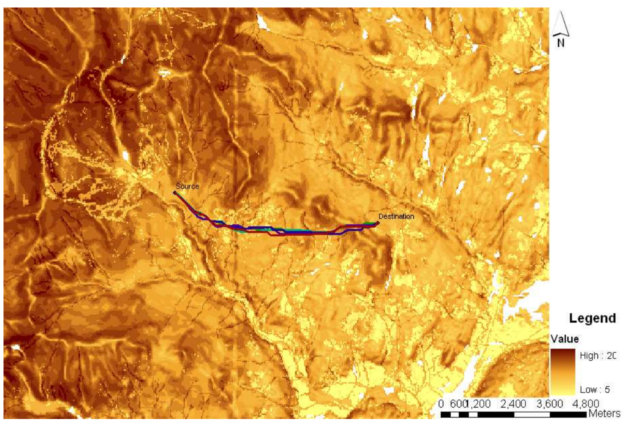

Fig. 8: The stability of repeated simulations using forward moving (3*3 neighborhood pattern 8-connectivity)

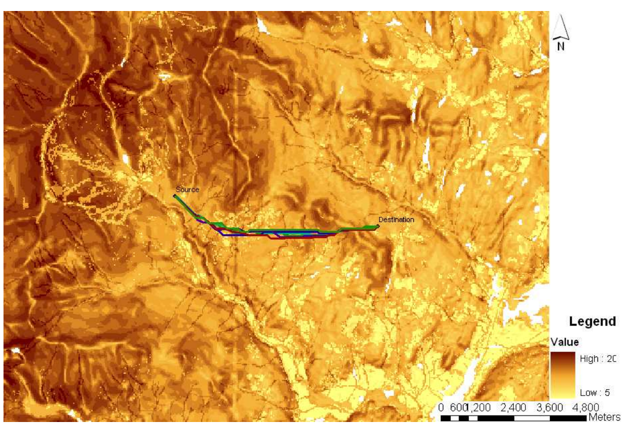

Fig. 9: The stability of repeated simulations using backward moving (3*3 neighborhood pattern 8-connectivity)

Table 3: The total cost distance of the paths, and the time (minute:second.millisecond) needed to calculate them.

Neighborhood Pattern	Dijkstra's algorithm		ABC algorithm	
	Cost	Time	Cost	Time
3*3	2247.45	27:46.81	2286.04	05:24.62
5*5	2178.44	28:51.34	2198.28	06:45.98
7*7	2082.45	33:46.70	2133.84	08:52.64

the use of large neighborhoods increases the computational expense of both the conventional Dijkstra's algorithm and the proposed ABC algorithm, it turns out that even the use of the largest neighborhood tested here entails an extra expense of about 20-60% compared to the use of the smallest neighborhood. For the proposed ABC algorithm, this increase in the processing time can be considered to be low when compared to the overall computational expense of the conventional Dijkstra's algorithm. The processing time of the conventional Dijkstra's algorithm is generally more than four times longer than the proposed ABC algorithm with different neighborhood patterns.

4 Conclusions

This paper presents a modified ABC algorithm for solving least-cost path problem in a raster-based GIS taking in to consideration how the quality of the raster-based paths can be improved by using larger connectivity patterns. Conventional ABC cannot complete a path on raster surfaces because of the lack of vision capability. Therefore, conventional ABC algorithm has been modified for adapting to the least cost path problem by incorporating a more sophisticated direction guidance search method to guide a bee walking toward the final destination more efficiently. The experimental results have shown that the performance of the modified ABC model is quite close to Dijkstra's algorithm while its

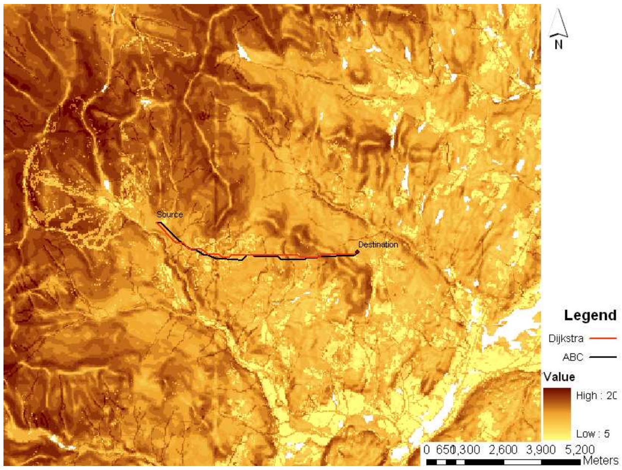

Fig. 10: Least-cost paths identified by the two algorithms (3*3 neighborhood pattern 8-connectivity)

computational complexity and solution time is much lower than Dijkstra's algorithm. The results also, indicate that raster-based paths can be improved by using larger connectivity patterns. Moreover, the proposed method can produce quite stable results from repeated simulations and there are also no significant changes of optimization results if the movement direction is reversed from forward moving to backward moving.

References

[1] C. Yu, J. Lee, M.J. Munro-Stasiuk: Extensions to least-cost path algorithms for roadway planning, International Journal of Geographical Information Science, **17**, 361-376 (2003).

[2] A.B. Gonalves: An extension of GIS-based least-cost path modelling to the location of wide paths, International Journal of Geographical Information Science, **24**, 983-996 (2010).

[3] H. Antikainen: Comparison of different strategies for determining raster-based least-cost paths with a minimum amount of distortion, Transactions in GIS, **17**, 96-108 (2013).

[4] T. Balstrom: On identifying the most time-saving walking route in a trackless mountainous terrain, Geografisk Tidsskrift, **102**, 51-58 (2002).

[5] W.G. Rees: Least-cost paths in mountainous terrain, Computers and Geosciences, **30**, 203-209 (2004).

[6] W. Collischonn, J.V. Pilar: A direction dependent least-cost-path algorithm for roads and canals, International Journal of Geographical Information Science, **14**, 397-406 (2000).

[7] W.N. Xiang: A GIS based method for trail alignment planning, Landscape and Urban Planning, **35**, 11-23 (1996).

[8] S. Bagli, D. Geneletti, F. Orsi: Routeing of power lines through least-cost path analysis and multicriteria evaluation to minimise environmental impacts, Environmental Impact Assessment Review, **31**, 234-239 (2011)

[9] S.C. Feldman, R.E. Pelletier, E. Walser, J.C. Smoot, D. Ahl: A prototype for pipeline routing using remotely sensed data and geographic information system analysis, Remote Sensing of Environment, **53**, 123-31 (1995).

[10] H. Lundqvist: Ecological cost-benefit modelling of herbivore habitat quality degradation due to range fragmentation, Transactions in GIS, **11**, 745-763 (2007).

[11] M.C.L Howey: Using multi-criteria cost surface analysis to explore past regional landscapes: A case studyof ritual activity and social interaction in Michigan, AD 1200-1600, Journal of Archaeological Science, **34**, 1830-1846 (2007).

[12] G. Davies, D. Whyatt: A least-cost approach to personal exposure reduction, Transactions in GIS, **13**, 229-246 (2009).

[13] P. Albert: Finding Minimal Cost Paths in Raster Geographic Information System Map Representations, Genetic Algorithms, Simulated Annealing and Tabu Search, Ph.D. Dissertation Kent State University College of Business, 137 (2004).

[14] M.F. Goodchild: An evaluation of lattice solutions to the problem of corridor location, Environment and Planning A, **9**, 727-38 (1977).

[15] J.V. Bemmelen, W. Quak, M.V. Hekken, P.V. Oosterom: Vector vs. raster-based algorithms for cross country movement planning, In Proceedings of the International Symposium on Computer-Assisted Cartography (Auto-Carto XI). Minneapolis, Minnesota, 304-317 (1993).

[16] H. Antikainen: Using the Hierarchical Path finding A* Algorithm in GIS to Find Paths through Rasters with Non uniform Traversal Cost, International Journal of Geo-Information, **2**, 996-1014 (2013).

[17] A. Sharma, V. Vyas, D. Deodhare: An Algorithm for Site Selection in GIS based on Swarm Intelligence, In 2006 IEEE Congress on Evolutionary Computation Sheraton Vancouver Wall Centre Hotel, Vancouver, BC: IEEE, 1020-1027 (2006).

[18] X. Li, J. He, X. Liu: Ant intelligence for solving optimal path covering problems with multi objectives, International Journal of Geographical Information Science, **23**, 839-857 (2009)

[19] D. Karaboga: An idea based on honey bee swarm for numerical optimization. Techn. Rep. TR06, Erciyes Univ. Press, Erciyes (2005).

[20] Z. Yin, X. Liu, Z. Wu: A Multiuser Detector Based on Artificial Bee Colony Algorithm for DS-UWB Systems, The Scientific World Journal, Article ID 547656, 1-8 (2013).

[21] D. Karaboga, B. Gorkemli, C. Ozturk, N. Karaboga: A comprehensive survey: artificial bee colony (ABC) algorithm and applications, Artif Intell Rev (2012).

[22] A.L. Bolaji, A.T. Khader, M.A. Al-betar, M. Awadallah: Artificial Bee Colony Algorithm, Its Variants and Applications: A Survey, Journal of Theoretical and Applied Information Technology, **47**, 434-459 (2013).

[23] A. Baykasoglu, L. Ozbakir, P. Tapkan: Artificial bee colony algorithm and its application to generalized assignment problem, Swarm Intelligence: Focus on Ant and Particle Swarm Optimization, 113-144 (2007).

[24] D. Karaboga, B. Basturk: On the performance of artificial bee colony (abc) algorithm. Applied Soft Computing, **8**, 687 - 697 (2008).

[25] ESRI, Spatial Analyst Tutorial, Redlands: ESRI Press.

Khalid A. Eldrandaly is an Associate Professor of Information Systems, Vice Dean for Academic Affairs, Faculty of Computers and Informatics, CIO and Director of the Communication and Information Technology Center, Zagazig University, Egypt. Dr. Eldrandaly is a certified GIS professional (GISP). He received his Ph.D. degree in Systems Engineering (GIS). He was a visiting scholar at Texas A&M University, USA, for two years. His research interests include GIS, Expert Systems, SDSS, MCDM, and Intelligent Techniques in Decision Making. Dr. Eldrandaly is a member of the Egyptian Engineers Syndicate, World Academy of Young Scientists (WAY), Arab Union of Scientists and Researchers (AUSR), Texas A&M International Faculty Network, Egyptian Software Engineers Association (ESEA), GIS Certification Institute, International Society for Environmental Information Science, and Association of Computing Machinery (ACM). He also currently serves as a member in the Review Committee of the IAJIT, IJGIS, IJOPCM, ASOC, ACIT, JEI, FCT and AJSE.

Mohammed M. Hassan an Associate Professor of Information Systems, Head of the Department of Information Systems, Faculty of Computers and Informatics, Zagazig University, Zagazig, Egypt. He is specialized in Computer Aided System Engineering. He received his PhD degree in Computer Aided System Engineering from the Department of System Engineering, Faculty of Engineering, Zagazig University, Zagazig, Egypt; also The Pennsylvania State University, State Collage, USA. He has published more than 28 articles in international journals and conference proceedings.

Nabil M. AbdelAziz is a Lecturer, Department of Information Systems, Faculty of Computers and Informatics, Zagazig University. He received his M.Sc. degree in Information Systems from Faculty of Computers and Informatics, Zagazig University. He is currently a Ph.D. student at the Department of Information Systems, Faculty of Computers and Informatics, Zagazig University. His research interests include GIS, Spatial Decision Support Systems, Multi-Criteria Decision Making, and Intelligent Techniques in Decision Making. He is a member of Egyptian Software Engineers Association (ESEA). He is the team leader of Quality Management Unit, Faculty of Computers and Informatics, Zagazig University.