A Novel Approach of Waves of Swarm with Case Based Reasoning To Detect Ground Water Potential

V.K. Panchal[1], Harish Kundra[2], Navpreet Kaur[3]

[1] *Defense Terrain Research labortary, DRDO, Delhi ;* vkpans@ieee.org
[2] *RIEIT, Railmajra ,Punjab ;* kundra_harish@yahoo.com
[3] *RIEIT, Railmajra, Punjab;* navpreet2k5@gmail.com

Abstract:
We have proposed a new groundwater possibility detection system based on particle swarm optimization (PSO) and case based reasoning (CBR). Groundwater constitutes an important source of water supply for various purposes, such as domestic industries and agriculture needs. Strategy is proposed based on natural process of evolution of organisms that are best adapted to the environment. A new concept of waves of swarm (WOS) derived from PSO is introduced in this study that operates on the problem case. Geological features (Geology, Landform, Soil Type, Land Use, Lineament, and Slope) and their corresponding solutions i.e. the possibility of ground water in fuzzy terms of high, moderate, or low are stored as cases in the case base. These features are selected as input parameters as they play a crucial role in detecting the groundwater potential. Each PSO particle is the intersection of multidimensional search space. We have assumed that each PSO particle is a set of geological features of groundwater cases. Rules are developed to refine the WOS with each iteration. The integration approach proposed here improves the retrieval accuracy of CBR using WOS.

Keywords:
Particle swarm optimization, Case based reasoning, Waves of swarm, Fitness function, Groundwater, K-Nearest Neighbor.

1. INTRODUCTION

1.1 Particle Swarm Optimization (PSO)

The particle swarm optimization (PSO) was originally designed by Kennedy and Eberhart [2]. Particle swarm optimization (PSO) is a population-based stochastic optimization technique modeled on the social behaviors observed in animals or insects, e.g., bird flocking, fish schooling, and animal herding. Since its inception, PSO has gained increasing popularity among researchers and practitioners as a robust and efficient technique for solving difficult optimization problems [3]. The technique involves simulating social behavior among individuals (particles) "flying" through a multidimensional search space, each particle representing a single intersection of all search dimensions [2]. In PSO, individual particles of a swarm represent potential solutions, which move through the problem search space seeking an optimal, or good enough, solution. The PSO approach utilizes a cooperative swarm of particles, where each particle represents a candidate solution, to explore the space of possible solutions to an optimization problem. Each particle is randomly initialized and then allowed to 'fly'. At each step of the optimization, each particle is allowed to evaluate its own fitness and the fitness of its neighboring

particles [6]. Each particle can keep track of its own solution, which resulted in the best fitness, as well as see the candidate solution for the best performing particle in its neighborhood.
At each optimization step, indexed by t, each particle, indexed by i, adjusts its candidate solution (flies) according to,

$$V_i(t+1) = V_i(t) + C1\ R1(X_{i,}p - Xi) + C2\ R2\ (X_{i,n} - X_i) \quad (1)$$
$$X_i(t+1) = X_i\ (t) + V_i\ (t+1) \quad (2)$$

Variables used in eq (1) and (2) are summarized below:

V_i = The Particle Velocity.
X_i = The Particle Position.
t= time.
R1 = A uniform random variable usually distributed over [0, 2].
R2 = A uniform random variable usually distributed over [0, 2].
$X_{i,p}$ = The particle's position (previous) that resulted in the best fitness so far.
$X_{i,n}$ = The neighborhood position that resulted in the best fitness so far.
C1 and C2 are acceleration constants.

Eq.(1) can be interpreted as follows. Particles combine information from their previous best position and their neighborhood best position to maximize the probability that they are moving toward a region of space that will result in a better fitness. The uniformly distributed random variables, R1, R2 are sampled for each i , t , and dimension of each vector X_i .

1.2. Case Based Reasoning

As a newly emerging artificial intelligence technology, the CBR provides a new approach to learn from the successful cases in the past and can imitate the human operation experiences [4].Case-based reasoning is one of the more recent developments in AI research and has now become an important and widely applied problem solving technology. It is based on the assumption that "similar problems have similar solutions" [5], In fact, this assumption is the guiding principle underlying most CBR systems. More precisely, the idea of CBR is to exploit the experience gained from similar problems in the past and to adapt then successful solutions to the current situation. In order to realize this idea, a CBR system has to maintain a structured memory of cases (also called a case base) which represents the experience and a means for specifying the similarity between cases. The basic notion of a case is thought of as a representation of knowledge about a specific situation or episode. Case-based reasoning is a well known and newly emerging artificial intelligence technique different from other reasoning approaches based on rule and on model. It has the advantages of simplified knowledge acquisition, high solution efficiency and easy knowledge accumulating [4]. It is widely used in many research fields such as disease diagnosis, market planning, engineering designing and so on. CBR reuses the past cases and experiences in a database called the case base to find a similar solution to the problem of a new situation. In a CBR system, the key factors that influence the retrieving of similar cases from case base include knowledge representation, attribute description and similarity measures definition [4]. Core problems addressed by CBR research can be

grouped into five areas. A set of coherent solutions to these problems constitutes a CBR method [1]:
* Knowledge representation
* Retrieval methods
* Reuse methods
* Revise methods
* Retain methods

2. WAVES OF SWARM

A general problem with optimization algorithms is that of becoming trapped in a local optimum. A particular algorithm may work well on one problem but may fail on another problem. Mathematical formulations for velocity are needed to adjust which makes it even more complicated. If an algorithm could be designed to adapt to the fitness function, adjusting itself to the fitness landscape, a more robust algorithm with wider applicability, without a need for problem specific engineering would result [6]. We propose a strategy based on natural selection i.e. "survival of the fittest" in which, when a particle fails to come close to the fitness function as compared to its competitor in previous swarm, that particle is simply discarded and particle which is more closer to the fitness function is allowed to survive. This is the type of search designed and analyzed in this paper. In PSO, the strategy of iteration and updating in PSO is that particles follow the best optimal particle in population. If a particle finds a local best position, the other particles fly rapidly toward it. May be these particles will trap in local optimum and cannot search in global resolution space again which is called premature convergence [4]. The results revealed that severe lack of population diversity will lead to premature convergence. Waves of swarm integrated with CBR are presented which is also based on procedure of dealing with population diversity. To design a better model of natural selection using the PSO algorithm, waves of swarm introduced in which new wave of swarm exist with each new iteration. In nature, individuals that have favorable features are more likely to boom and reproduce. The favorable adaptation is assumed to protract the lifetime of an individual whereas unfavorable adaptations reduce the lifespan of an individual or group. CBR is selected as one of the main modeling tool. The integration approach proposed here improves the retrieval accuracy of CBR using waves of swarm.

2.1 Natural selection in waves of swarm

The primary assumptions made to implement waves of swarm are:

- A particle in swarm will (be rewarded) have its lifetime extended by finding a more fit state as compared to its competitor in previous swarm.
- A particle in swarm will (be penalized) simply discarded for failing to find a more fit state as compared to its competitor.

These ideas implement an algorithm imitating natural selection.

3. GROUNDWATER POSSIBILITY DETECTION SYSTEM

"If the wars of the twentieth century were fought over oil, the wars of this century will be fought over water." [The World Bank]. Water is essential for life. Groundwater constitutes only 0.6% of all the water on this earth planet, 97.4% accounts for seawater and 2% for snow and ice on the poles [7]. Hence groundwater is an important commodity which we use for various purposes such as domestic, industrial and agricultural use but with the increase in population its resources are depleting and hence the necessity to find its resources arises. Here detection for possibility of ground water at a particular region is taken as an application.

3.1. Problem Formulation

Querying a large case-base of groundwater detection for past groundwater cases similar to present case using the waves of swarm k-nearest neighbor algorithm that is designed and tuned with the help of a groundwater detection expert, can increase the accuracy of predictions of presence of groundwater at a particular region.

3.2 Integration of waves of swarm with CBR

In a CBR system, the key factors that influence the retrieving of similar cases from case base include knowledge representation; attribute description and similarity measures definition [4]. In this CBR system, geographical parameters and their corresponding solutions, i.e. the possibility of ground water (High, Moderate, and Low) are stored as cases in the case base. By the attribute importance, six attributes are selected as the inputs. The possibility of ground water is selected as the output decision attribute of the case. The cases are collected from the expert knowledge. The case is represented by the ordered pair C= {F, D}, where F is the geographical parameter of case C and F = {Geology, Landform, Soil Type, Lineament, Slope, Land Type}, D is the decision attribute, i.e. the possibility of ground water in the given region.

3.3 Particle & Swarm initialization

Each PSO particle is a set of features, stored as a case in case base. Wave of swarm consisting of particles is randomly initialized with index of cases, x_i stored in CBR on an appropriate range $x_{min} <= x_i <= x_{max}$.

3.4 Similarity Criteria

The case searching and matching is a key step in case retrieving and it directly influences the retrieval efficiency and accuracy. In fact, the case retrieval in essence is to find the most similar case in the case base to target case. K-Nearest Neighbor (KNN) is widely used for its advantages of clear physical concept and simple calculation for case searching [4]. A sum of similarities is calculated according to the similarity between each feature in problem case and the cases in case base. Evaluate the fitness of each particle for case retrieval according to following equation:

$$\mathbf{Sim}_{(P,C)} = \sum_{a=1}^{n} f(P_a, C_a) * w_a \quad \left\{ \begin{array}{l} \text{if } P_a == C_a \\ f(P_a, C_a) = 1 \\ \text{otherwise} \\ f(P_a, C_a) = 0 \end{array} \right. \qquad (3)$$

Where, $Sim_{(P,C)}$ in eq(3) represents the similarity degree of problem case(P) and case(C) stored in case base, a= attribute of case, n=number of attributes, P_a= problem case, C_a=cases stored in case base and w_a=weight of attribute a, the more important attributes should assigned larger weights then less important ones. The larger the weighted sum is, the more similar the two cases will be. Only that case whose weighted sum is greater than the pre-specified value will be retrieved as similar case. The degree (percentage) for possibility of presence of ground water at a given area is calculated as weighted average output, as given below:

$$\mathbf{Out = (Sim_{(P,C)} / total\ Weight) * 100} \quad (4)$$

Where, $Sim_{(P,C)}$ represents the similarity degree of problem case (P), case (C) stored in case base and total weight is the sum total weight of all the six geographical attributes.

3.5. The Algorithm

Initially the fitness of each particle is calculated and stored. Number of waves is equal to number of iterations. In next iteration, when new wave arrives, fitness of each particle in new wave is evaluated and compared with the fitness of particle at the corresponding position in previous swarm (the competitor). If the fitness value of current particle is better than the best fitness value (pbest) of its competitor. Then, set current particle's value as the new **pbest.** The particle which is closer to fitness function is allowed to stay in swarm and its competitor is discarded. At the end of each iteration, choose the particle with the best fitness value of all the particles as the **gbest.** An analogy "survival of the fittest" is used here to form the refined set of particles in a swarm with each iteration. This process continues until fitness function is fully satisfied or specified number of iterations is completed. This is shown in pseudo code below:

```
Main Program Loop
For each wave of swarm in the collection
     For each particle in the swarm
          Initialize each particle randomly with
          cases stored in case base.

     If particle finds more fit state
     than competitor
          Reward particle:  Extend particle life

     If particle has not improved
          Penalize particle: Discard particle:
          Reduce particle life

Update wave with fit Particle: pbest

At the end of each iteration:
      Choose particle with best fitness as the gbest
```

Figure 1. Pseudo code for WOS

4. RESULTS

The previous section describes the Groundwater detection system, its components, and how it is configured. The algorithm for groundwater detection system using waves of swarm (shown in Figure 1) is coded in Matlab 7.0.

4.1 Dataset

The dataset used in our experiment consists of groundwater observations made by domain experts to predict the groundwater possibility.

Table 1: Six attributes used in dataset

Attributes	Values
Geology	Sedimentary, Younger alluvium, Older alluvium, Igneous,Metamorphic
Landform	Floodplain,Intermontanevalley,Pediment,Alluvialfans,Bajada,Pediplain, Buriedpediment,AlluvialPlain,DeltaicPlain,Wadi,Riverterraces,Oldmeander etc.
Soil	Sandyloam,Sandygravel,Coarsesand,Clayloam,Alluvialsand,Gravelsand, GravelSandPebbles,Sand,Rocky etc.
Land use	Agriculturalland,Forest,Cultivatedland,Fallowland,Waterbody,Wasteland, Swampy land, Buildup, Urban ,Grass, Shrubs ,mixed vegetation etc.
Slope	Gentle, Steep
Lineament	Absent, Present

First column of table 1 gives the categorical attributes that are used in the dataset and the second column indicates the attribute values. Results of system are presented in the form of graphical user interface, graphs, and interpreted.

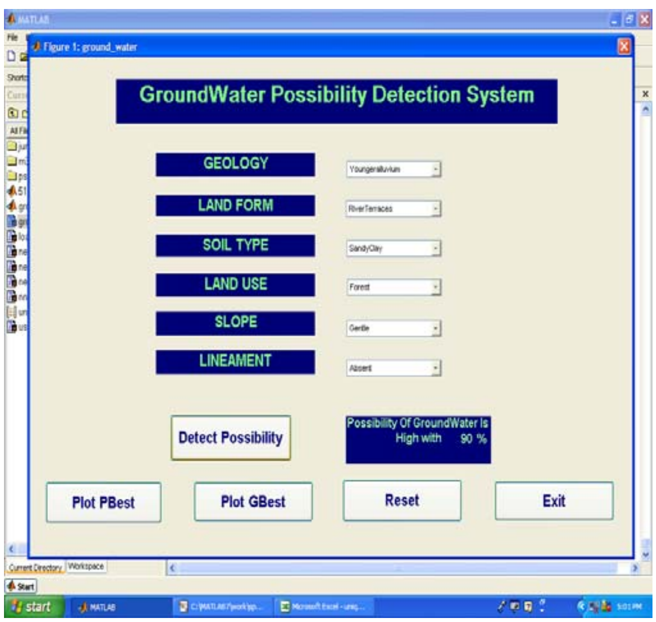

Figure 2. Graphical User Interface for Groundwater possibility detection system

Input given to six parameters is shown below:

Table 2: input values

Geology	Younger alluvium
Landform	River Terraces
Soil type	Sandy Clay
Land Use	Forest
Slope	Gentle
Lineament	Absent

Graph in figure 3. Plots the pbest values of particles (represented by '*'). X-axis represents the number of iterations where:

Maximum number of iterations = (size of Casebase / number of particles) (5)

Y-axis represents the pbest values of particles in each wave.

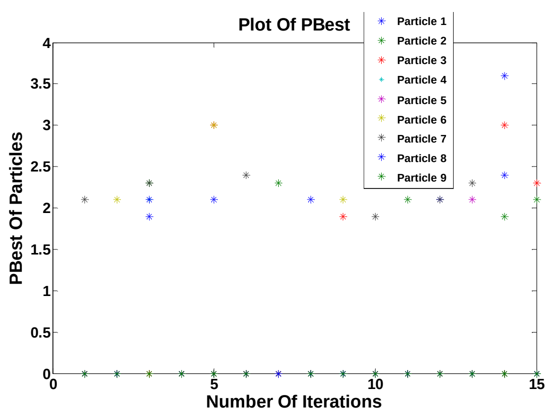

Figure 3. Plot of PBest for above case

Figure 4. Plot of GBest for above case

X-axis in Graph of figure 4. Represents number of iterations and y-axis shows global best positions in each iteration (represented by '*'). At the end of each iteration the particle with the best fitness value of all the particles is chosen as the **gbest.**
Output of the input combinations of values is shown in text box in fig 2:

"Possibility of groundwater is High with 90%".

5. CONCLUSION AND FUTURE SCOPE

We have proposed, implemented, and tested the WOS with case based reasoning based detection system called groundwater detection system. The detection for presence of ground water at a particular region is the increasing need of today's time. The proposed methodology can be used to predict the presence of groundwater at inaccessible areas through remote sensing. We proposed WOS to predict this possibility, overcoming the difficulties and disadvantages of the traditional methods. Integration of waves of swarm based on PSO with case based reasoning is proposed to improve the retrieval accuracy in weighted KNN case retrieval. Further research is needed to explore the appropriate configuration and usage of waves of swarm. Proposed work can be explored with different kinds of case bases, multiple case bases simultaneously can be investigated. Other possible applications of this extended version of particle swarm optimization technique are processing medical data, where we need to diagnose the disease of patients. Other areas such as oceanographic astronomical observations can be attacked.

7. REFERENCES

[1] Agnar Aamodt, Enric Plaza, "Case Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches". AICom Artificial Intelligence Communications, IOS Press, Vol 7:1, pp 39-59, 1994.

[2] Anthony Carlisle, Gerry Dozier, "Adapting particle swarm optimization to Dynamic Environments". Proceedings of international conference on artificial intelligence, Las Vegas, USA, pp 429-434, 2000.

[3] Christian Blum and Daniel Merkle, "Swarm Intelligence Introduction and Applications". Springer, pp 43, 58-60, 2008.

[4] Chunhua Yang, Hongqiu Zhu, Weihua Gui, "Permeability prediction model for imperial smelting furnace based on improved case-based reasoning", IEEE Proceedings of the 7th world congress on intelligent control and automation, June 25-27, 2008, Chongqing, China.

[5] Eyke Hullermeier,"Case based approximate reasoning", Springer, pp 30-32, 2007.

[6] Jason Tillett, T.M Rao, Ferat Sahin and Raghuveer Rao, "Darwinian particle swarm optimization", Proceedings of 2nd Indian international conference on Artificial intelligence, December 2005.

[7] I.Pulford, H.Flowers, "Environmental chemistry at a glance", Great Britain: Blackwell, pp 46-47, 2006.