Applying Neural Network Ensembles for Clustering of GPS Satellites

Mosavi, M. R.,[1] and Azami, H.,[2]
Department of Electrical Engineering, Iran University of Science and Technology, Narmak, Tehran 16846-13114, Iran, E-mail: M_Mosavi@iust.ac.ir (Corresponding Author)[1], azamihamed@yahoo.com[2]

Abstract
Several external sources introduce errors into a GPS position estimated by a GPS receiver. One important factor in determining positional accuracy is the constellation, or geometry, of the group of satellites from which signals are being received. Geometric Dilution of Precision (GDOP) allows to determine the good selected satellite geometry and to give an approach about the GPS accuracy. The most correct solution to get GPS GDOP is to use inverse matrix on all the combinations and selecting the lowest one, but inversing a matrix puts a lot of computational burden on the navigator's processor. In this paper, the clustering of GPS satellites is done based on the GDOP resulting from the Neural Network (NN), NN ensemble and NN ensemble by using Genetic Algorithm (GA) to select the appropriate subset of navigation satellites. The proposed methods are simulated and validated by a software simulation. These algorithms provide a realistic computational approach without needing to calculate the inverse matrix. The simulation results demonstrate that NN ensemble with GA has greater accuracy and can improve clustering accuracy of GPS satellites about 99%.

1. Introduction

The Global Positioning System (GPS) accuracy depends mainly on the geometry of the satellites. To check the accuracy of the position, the user equivalent range error is multiplied by a factor which depends on the geometry of satellite's constellation. This factor is the Geometric Dilution of Precision (GDOP) (Yarlagadda et al., 2000). DOP only depends on the position of the satellites. The computed position can vary depending on which satellites are used for the measurement. Different satellite geometries can magnify or lessen the position error. A greater angle between the satellites of lowers the DOP provides a better measurement. A higher DOP indicates poor satellite geometry and an inferior measurement configuration. Table 1 shows DOP ratings.

Table 1: DOP ratings

DOP value	Ratings
1	Ideal
2-3	Excellent
4-6	Good
7-8	Moderate
9-20	Fair
21-50	Poor

The most correct solution to calculate GPS GDOP is to use inverse matrix on all the combinations of GPS satellites and selecting the lowest one, but inversing a matrix consumes a lot of computational burden on the navigator's processor. Jwo and Lai, 2007 used Neural Network (NN) to reduce this computational burden. A NN is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neuron) working in unison to solve specific problems. The NN, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained NN can be thought of as an expert in the category of information (Mosavi, 2007a). Clustering is finding groups of objects such that the objects in a group will be similar (or related) to one another and different from (or unrelated to) the objects in other groups. Ensemble technique is a powerful technique for clustering. Usually, the combinational methods have more accuracy, power, and generality rather than single clustering (Parvin et al., 2009). Ensemble methods usually cause in the improvement of clustering, because classifiers with different characterizes and methodologies can complete each other (Webb and Zheng, 2004). Condorcet Jury theorem said that combination of

classifiers operates usually better than single classifier. It indicates if more diverse classifiers are used in the ensemble, so error of them can significantly be reduced. For enhancing the performance of NN ensemble, Mohammadi at al., 2008 and Chen and Yao, 2008 use Genetic Algorithm (GA). The GA is robust and intelligent algorithm for performing a random search of a defined N-dimensional solution space. The GA does not require any gradient information and therefore may be suitable for non-linear problems. The GA can enhance performance of clustering ensemble. In this paper, NN ensembles are proposed for the clustering of GPS satellites. This paper is organized as follows. In section 2, the concept of GPS GDOP is briefly reviewed. The NN, NN ensemble and NN ensemble by using GA are explained for clustering of GPS satellites in sections 3 and 4. Section 5 provides computer simulations and discussions. Conclusions are given in the section 6.

2. The Concept of GPS GDOP

Basically, the GPS accuracy is relied to the GDOP. Figure 1 shows the geometry of the satellites and its affect on the GDOP values. There are usually more satellites available than a receiver needs to fix a position, so the receiver picks a few and ignores the rest. If it picks the satellites that are close together in the sky (Figure 1a), the intersecting circles that define a position will cross at very shallow angles. That increases the error margin around a position. If it picks satellites that are widely separated (Figure 1b), the circles intersect at almost at right angle and that minimizes the error region. Good DOP, a small value representing a large unit-vector-volume, results when angles from receiver to satellites are different. Where as poor DOP, a large value representing a small unit vector-volume, results when angles from receiver to the set of satellites used are similar. We resume the definitions of GDOP computation, useful for the sequel of the document and help understanding our contributions. The absolute distance between a user and a satellite is defined as follows (Zirari et al., 2009):

$$R_i = \rho_i + \Delta\rho^{iono,i} + \Delta\rho^{trop,i}$$

Equation 1

where:

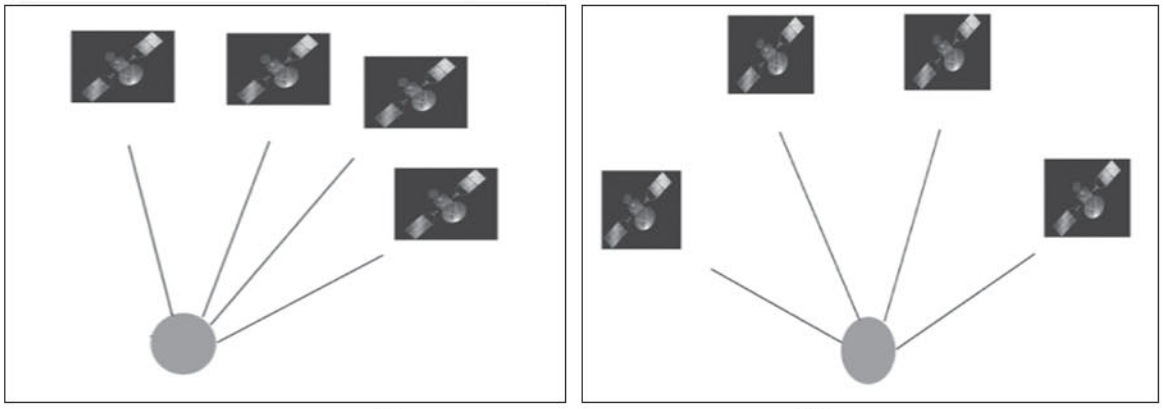
$$\rho_i = \sqrt{(X_i - X_u)^2 + (Y_i - Y_u)^2 + (Z_i - Z_u)^2} - \delta t_u$$

Equation 2

$\Delta\rho^{iono,i}$ and $\Delta\rho^{trop,i}$ which are the errors induced by the ionospheric and the tropospheric propagation, are calculated from a model, $(X_u, Y_u, Z_u, \delta t_u)$ are the four system unknowns and δt_u is the correction the receiver has to apply to its own clock. To resolve this system, we need four equations which mean four pseudo-ranges from four different satellites. The pseudo-ranges can be approximated by a Taylor expansion. We obtain:

$$\hat{\rho}_i = \sqrt{(X_i - \hat{X}_u)^2 + (Y_i - \hat{Y}_u)^2 + (Z_i - \hat{Z}_u)^2} - \delta \hat{t}_u$$

Equation 3



(a) (b)

Figure 1: Satellite's diagram and its relation with DOP: (a) Bad DOP and (b) Good DOP

The Taylor expansion at the first order is:

$$\Delta\rho_i = \rho_i - \hat{\rho}_i = a_{xi}\Delta x_u + a_{yi}\Delta y_u + a_{zi}\Delta z_u - c\Delta t_u$$

Equation4

where:

$$a_{xi}=\frac{X_i-\hat{X}_u}{\hat{r}_i};\ a_{yi}=\frac{Y_i-\hat{Y}_u}{\hat{r}_i};\ a_{zi}=\frac{Z_i-\hat{Z}_u}{\hat{r}_i};\ \hat{r}_i=\sqrt{(X_i-\hat{X}_u)^2+(Y_i-\hat{Y}_u)^2+(Z_i-\hat{Z}_u)^2}$$

Equation 5

Let assume N_{sat} be the number of visible satellites. The matrix H is as follows:

$$H=\begin{bmatrix} a_{x1} & a_{y1} & a_{z1} & 1 \\ a_{x2} & a_{y2} & a_{z2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{xN_{sat}} & a_{yN_{sat}} & a_{zN_{sat}} & 1 \end{bmatrix}$$

Equation 6

Let define the G matrix:

$$G=(H^T H)^{-1}$$

Equation 7

The GDOP is:

$$GDOP=\sqrt{trace[G]}$$

Equation 8

3. Artificial Neural Network

An appropriate model of the nervous system should be able to produce similar responses and behaviours in artificial systems. The nervous system is built by relatively simple units, the neurons, so copying their behaviour and functionality should be the solution. A NN is composed of many artificial neurons that are linked together according to specific network architecture. The objective of the NN is to transform the inputs into meaningful outputs. Figure 2 shows a NN with 3 layers, first layer include p neurons as input, second layer have q neurons as hidden layer and 3rd layer have one neuron as output. Neurons work by processing information.

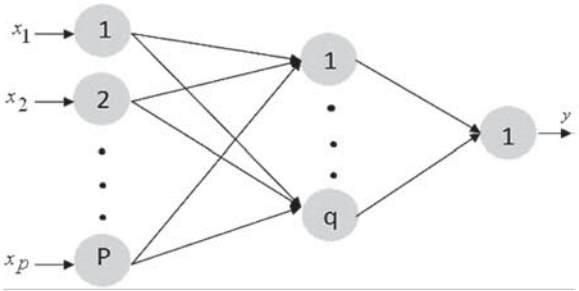


Figure 2: NN architecture

They receive and provide information in form of spikes:

$$y_k = f(\sum_{l=1}^{m} w_{kl}x_l)$$

Equation 9

where y is the neuron's output, X is the vector of inputs, W is the vector of synaptic weights and f is node activation function, typically the sigmoid:

$$y=\frac{1}{1+e^{-ax}}$$

Equation 10

3.1 NN Learning using Back Propagation

In order to train a NN, the NN weights can be adjusted so that the errors between the desired output and actual output are reduced. The Back Propagation (BP) algorithm is the most widely used method for determining the NN weights. Totally the BP has two phases: forward and backward. The forward pass phase computes feed-forward propagation of input pattern signals through network. For hidden units (Mosavi, 2007b):

$$z_j(t) = f(\sum_{i=1}^{p} w_{ji}(t)x_i(t))$$

Equation 11

For output units:

$$y(t) = f(\sum_{j=1}^{q} w_j(t)z_j(t))$$

Equation 12

The backward pass phase propagates the error signal backwards through network starting at output units (where the error is the difference between actual and desired output values). Cost function E is defined as sum of measure error squares:

$$E(t) = \frac{1}{2}e(t)^2 = \frac{1}{2}[y(t) - d(t)]^2$$

Equation 13

where d and y are the target and output values, respectively. We want to know how to modify weights in order to decrease E. We use gradient descent, i.e.:

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t) = w_{ji}(t) - \eta \frac{\partial E(t)}{\partial w_{ji}(t)}$$

Equation 14

where the $\eta > 0$ is learning factor. Both for hidden units and output units, the partial derivative can be rewritten as product of three terms using chain rule for partial differentiation:

$$\frac{\partial E(t)}{\partial w_{ji}(t)} = \frac{\partial E(t)}{\partial y(t)} \frac{\partial y(t)}{\partial z_j(t)} \frac{\partial z_j(t)}{\partial w_{ji}(t)}$$

Equation 15

These terms are:

$$\frac{\partial E(t)}{\partial y(t)} = e(t),\ \frac{\partial y(t)}{\partial z_j(t)} = w_j(t) f'\left(\sum_{j=1}^{q} w_j(t) z_j(t)\right),$$

$$\frac{\partial z_j(t)}{\partial w_{ji}(t)} = x_i(t) f'\left(\sum_{i=1}^{p} w_{ji}(t) x_i(t)\right)$$

Equation 16

In fact in every stage, we update $\Delta w_{ji}(t)$ until this parameter achieve to its desire. For $w_j(t)$ and the bias vectors, we can use the same method too.

3.2 NN Learning using Clustering Ensemble

Ensemble methods construct a set of classifiers from the training data and predict the classes of test instances by combining the predictions of these classifiers. Clustering ensemble combines some clustering outputs to obtain better results.

High robustness, accuracy and stability are the most important characteristics of clustering ensembles (Topchy et al., 2004). To keep away from the sample subset too simple, the sample subsets are interpolated some samples which are randomly selected from the whole sample space. By contrast experiments, the recent learning algorithm is proved that can decrease the prediction error of NN ensemble and improves the prediction accuracy (Zhi-Gang and Nan-Feng, 2009). There are two main ways for clustering ensemble of data, namely, bagging and boosting. There are several examples of these techniques in (Zhi-Gang and Nan-Feng, 2009, Dudoit and Fridlyand, 2003 and Freun and Schapire, 1995). Constructing an ensemble of classifiers involves the following steps (Masson and Denoeux, 2009), as shown in Figure 3.

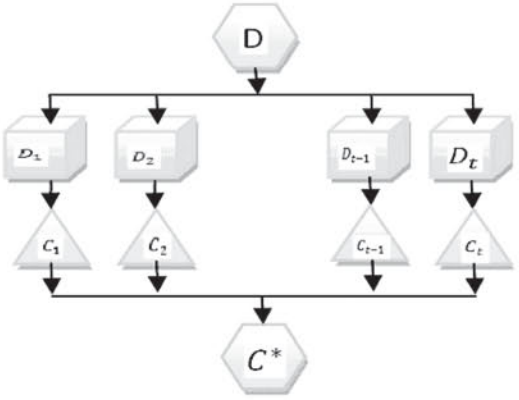


Figure 3: An illustrative example of the ensemble learning method (Parvin et al., 2009)

Step 1: In this step, we try to create different subset of data $D1, D2, \cdots, Dt$ in order to make independent classifiers. Each subset may be exact duplicates or a slightly modified version of the original data. In most cases, the size of each subset is identical to the original size, although class training sample distribution may have changed.

Step 2: We use subset of features for obtaining diversity in ensemble. In this method each classifier is trained on different subset of features.

Step 3: Classify the test instance X by using combining the predictions made by the base classifiers:

$C(X) = Vote[C1(X), C2(X), \cdots, Ct(X)]$ .The voting scheme may simply be a majority vote of the individual predictions or weighted by the accuracy of the base classifiers. The NN is retrained with different sets of data.

Creating a number of new train sets is executed by clustering ensemble technique. In this method, some of train samples are iterated more than ones. Then, the new train sets are created by appearing each of these clusters equal or more than one time.

3.3 NN Ensemble Learning using Genetic Algorithm

The GA is an optimization and search technique based on the principles of genetics and natural selection. The GA has many benefits, for example, its concept is easy to understand, supports multi-objective optimization, good for noisy environments, answer gets better with time, inherently parallel and easily distributed. A GA has three major components. The first component is related with the creation of an initial population of m randomly selected individuals. The initial population shapes the first generation. The second component inputs m individuals and gives as output an evaluation for each of them based on an objective function known as fitness function. This evaluation describes closeness value of our demands respect to each one of these m individuals. Finally, the third component is responsible for the formulation of the next generation. A new generation is formed based on the fittest individuals of the previous one. This evaluation procedure of generation N and production of generation $N+1$ (based on N) is iterated until a performance criterion is met. The creation of offspring based on the fittest individuals of the previous generation is known as breeding. The breeding procedure includes three basic genetic operations: reproduction, crossover and mutation. Reproduction selects probabilistically one of the fittest individuals of generation N and passes it to generation $N+1$ without applying any changes to it. On the other hand, crossover selects probabilistically two of fittest individuals of generation N; then in a random way chooses a number of their characteristics and exchanges them in a way that the chosen characteristics of the first individual would be obtained by the second and vice versa. Following this procedure creates two new offsprings that both belong to the new generation. Finally, the mutation selects probabilistically one of the fittest individuals and changes a number of its characteristics in a random way. The offspring that comes out of this transformation is passed to the next generation (Ru and Jianhua, 2008). The GA with large capability of searching in very big spaces can be applied for finding the best sets, efficiently. Each chromosome in GA demonstrates a new set, so that the i-th gene determines the number of times that the i-th cluster should be appeared in the new set.

The population of the last generation is considered to be as the final modified train sets.

4. GPS GDOP Clustering by Neural Network

When making the instructional data, all of the input and output variants between 0 and 1 will be normalized to reduce the instruction time. Since $H^T H$ is a 4×4 matrix, it has four $\lambda_i (i=1,2,3,4)$ eigenvalue. We know that the four eigenvalue for the matrix $(H^T H)^{-1}$ will be λ_i^{-1}. Based on the fact that the trace of a matrix is the sum of its eigenvalues, equation (8) would be as below (Jwo and Lai, 2007 and Doong, 2009):

$$GDOP = \sqrt{\lambda_1^{-1} + \lambda_2^{-1} + \lambda_3^{-1} + \lambda_4^{-1}}$$

Equation 17

Mapping with the definition of four variants would be done as below:

$$x_1(\vec{\lambda}) = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = trace(H^T H)$$

Equation 18

$$x_2(\vec{\lambda}) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 = trace[(H^T H)^2]$$

Equation 19

$$x_3(\vec{\lambda}) = \lambda_1^3 + \lambda_2^3 + \lambda_3^3 + \lambda_4^3 = trace[(H^T H)^3]$$

Equation 20

$$x_4(\vec{\lambda}) = \lambda_1\lambda_2\lambda_3\lambda_4 = \det(H^T H)$$

Equation 21

The GDOP can be considered as a mapping directly from $\vec{x}$ to GDOP classes. Mapping from $\vec{x}$ to the GDOP classes is extremely non-linear and cannot be determined analytically, but it can be determined precisely by NN. In this paper, NN is designed to do the mapping from $\vec{x}$ to GDOP classes. Figure 4 shows the total clustering block diagram of GPS GDOP using NN.

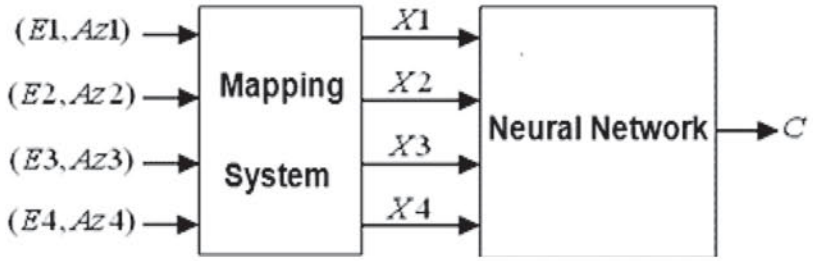


Figure 4: Clustering block diagram of GPS GDOP using NN

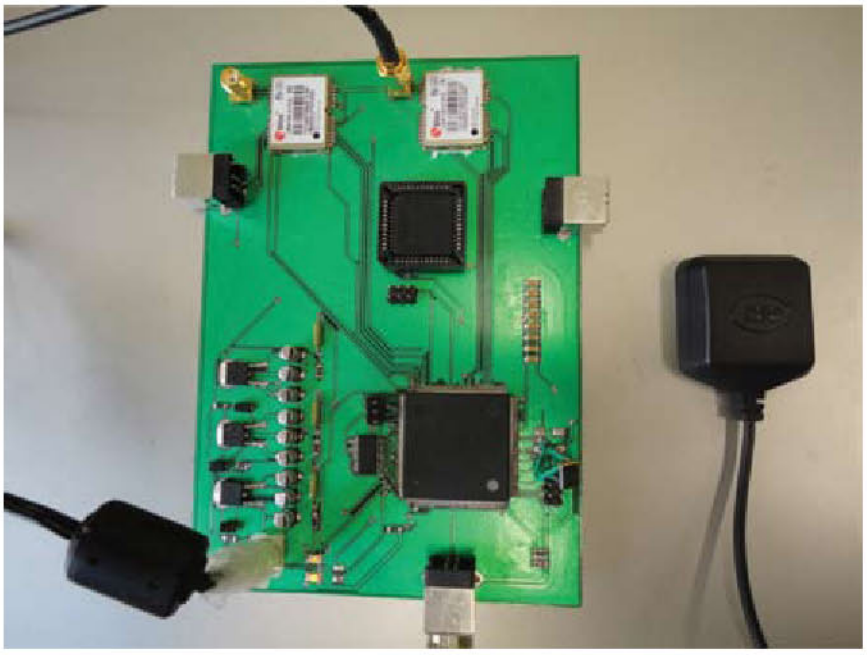


Figure 5: GPS data collecting embedded system

5. Simulations and Discussions

In this research, a large set of experiments was carried out using the following set-up: a standard GPS receiver was installed in a fixed point and was connected to a PC. In order to GPS GDOP real collection, the azimuth and the elevation of each observed satellite are measured by using a developed embedded system. After collecting the GPS information on DRAM, these data were transformed to serial port of PC for processing. Figure5 shows the entire GPS data collecting embedded system. All MLPs were one hidden layer, with the difference neurons between 500 and 5000. The sigmoid function was applied in this paper for training MLPs. The GA applied here had 20 populations and the number of iterations was 20. The crossover and mutation functions were assumed to be one-points and Gaussian, respectively. Three classes or outputs for the proposed NNs are considered for GPS GDOP clustering. When the GDOP is higher than the first threshold, NN outputs are [1 0 0] (for the large state) and when the GDOP is placed between the first and second threshold, NN outputs are [0 1 0] (for the average case) and when the GDOP is smaller than the second threshold, NN outputs are [0 0 1] (for the low state). Similarly, clustering can be done more than three groups. Tables 2 to 4 represent accuracy of GPS GDOP clustering, iterations and number of hidden layers neurons for three proposed methods.

Table 2: Accuracy of GPS GDOP clustering using NN

Number of neurons	500 iterations	1000 iterations	2000 iterations	3000 iterations	4000 iterations	5000 iterations
5	88.2	88.3	88.5	89	90.2	94.5
10	90	91	92.4	91.9	92.6	94.7
20	91	93.4	95.8	95	96	97.1
30	92	95	97.2	97.7	97	97
40	92.4	96	96.4	97	97	97.5
50	94.4	96	96.5	97.7	97	97

Table 3: Accuracy of GPS GDOP clustering using NN ensemble

Number of neurons	500 iterations	1000 iterations	2000 iterations	3000 iterations	4000 iterations	5000 iterations
5	89.1	88.8	89	89.1	90.6	95.6
10	90	91.2	93.3	92.2	92.8	95.9
20	91	93.8	96.8	95.5	96.6	97.6
30	92.6	96	97.7	97.7	97.9	97.7
40	93.4	97.1	97	97.2	97.3	97.6
50	94.1	97.4	97	98.3	97.3	97.7

Table 4: Accuracy of GPS GDOP clustering using NN ensemble with GA

Number of neurons	500 iterations	1000 iterations	2000 iterations	3000 iterations	4000 iterations	5000 iterations
5	89.5	89.5	89.5	89.5	91	98.5
10	91	93	94.7	95	93	98.5
20	92.5	96.2	97.7	96.5	98	98.7
30	94.7	97	97.9	97.7	97	98.5
40	97.5	97.8	98	98	97.7	99
50	97.8	97	98.1	98.1	97.6	98.9

Table 5: Comparison of accuracy of the proposed methods

Methods	Max	Min	Mean
NN	97.7	88.2	94.17
NN ensemble	98.3	88.8	94.71
NN ensemble using GA	99	89.5	95.85

Table 5 shows comparing accuracy between these three methods for GPS GDOP. The simulation results demonstrate that GPS GDOP clustering using NN ensemble with GA approach has better accuracy than other approaches. Only Jwo and Lai, 2007 used NN for GPS GDOP clustering with BP technique for NN learning. Advantages of the proposed method based on NN ensembles in our paper than reference of mentioned are high accuracy and low CPU time. It has the structure complexity less than for hardware implementation. It also requires the memory less than for software implementation.

6. Conclusion

The GDOP represents the geometrical relationship between positioning accuracy and satellite configuration. It is calculated using the positions of four visible satellites and the position of the receiver. The ideal DOP value is one. Values of two to three are considered excellent, while values greater than nine are considered high and should be discarded. Existing methods for GPS GDOP approximation and clustering are time-consuming. Intelligent tools exist to assess the ideal satellite configuration or DOP value at a given time.

In this paper, we propose three methods for GPS GDOP clustering using NN, NN ensemble and NN ensemble with GA. In the first method, we use NN with BP learning algorithm. The performance of classifier was increased by combining of classifiers in second method. In third method, we use GA for enhancing the ensemble clustering. The simulation results demonstrate that GPS GDOP clustering using NN ensemble with GA have better accuracy and better CPU time than simple clustering using NN and also NN ensemble.

References

Chen, M., and Yao, Z., 2008, Classification Techniques of Neural Networks using Improved Genetic Algorithms, *IEEE Conference on Genetic and Evolutionary Computing*, 115-119.

Dudoit, S., and Fridlyand, J., 2003, Bagging to Improve the Accuracy of a Clustering Procedure, *Oxford Journals of Life Sciences Mathematics and Physical Sciences, Bioinformatics*, Vol.19, No.9, 1090-1099.

Doong, S. H., 2009, A Closed-Form Formula for GPS GDOP Computation, *Journal of GPS Solutions,* Vol.13, No.3, 183-190.

Freun, Y., and Schapire, R., 1995, A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting, *Second European Conference on Computational Learning Theory*, 23-37.

Jwo, D. J., and Lai, C. C., 2007, Neural Network-based GPS GDOP Approximation and Classification, *Journal of GPS Solutions*, Vol.11, No.1, 51-60.

Masson, M. H., and Denoeux, T., 2009, Belief Functions and Cluster Ensembles, *Lecture Notes in Computer Science*, Vol.5590, 323-334.

Mohammadi, M., Alizadeh, H., and Minaei-Bidgoli, B., 2008, Neural Network Ensembles using Clustering Ensemble and Genetic Algorithm, *IEEE Conference on Convergence and Hybrid Information Technology*, 761-766.

Mosavi, M. R., 2007a, Precise Real-Time Positioning with a Low Cost GPS Engine using Neural Networks, *Journal of Survey Review*, Vol.39, No.306, 316-327.

Mosavi, M. R., 2007b, GPS Receivers Timing Data Processing using Neural Networks: Optimal Estimation and Errors Modeling, *Journal of Neural Systems*, Vol.17, No.5, 383-393.

Parvin, H., Alizadeh, H., and Minaei-Bidgoli, B., 2009, Using Clustering for Generating Diversity in Classifier Ensemble, *Journal of Digital Content Technology and Its Applications*, Vol.3, No.1, 51-57.

Ru, N., and Jianhua, Y., 2008, A GA and Particle Swarm Optimization based Hybrid Algorithm, *IEEE Congress on Computational Intelligence*, 1047-1050, 2008.

Topchy, A., Minaei-Bidgoli, B., Jain, A. K., and Punch, W. F., 2004, Adaptive Clustering Ensembles, *IEEE Conference on Pattern Recognition*, Vol.1, 272-275.

Webb, G. I., and Zheng, Z., 2004, Multistrategy Ensemble Learning: Reducing Error by Combining Ensemble Learning Techniques, *IEEE Transaction on Knowledge and Data Engineering*, Vol.16, No.8, 980-991.

Yarlagadda, R., Ali, I., Al-Dhahir, N., and Hershey, J., 2000, GPS GDOP Metric, *IEE Proc.-Radar, Sonar Navig.*, Vol.147, No.5, 259-264, .

Zhi-Gang, G., and Nan-Feng, X., 2009, A New Ensemble Learning Algorithm Based on Improved K-Means for Training Neural Network Ensembles, *Second International Symposium on Intelligent Information Technology and Security Informatics*, 8-11.

Zirari, S., Canalda, P., and Spies, F., 2009, Modelling and Emulation of an Extended GDOP for Hybrid and Combined Positioning System, *European Navigation Conference-Global Navigation Satellite Systems*, Italy, May.