An Energy Aware Invasive Weed Optimization Based Clustering in Wireless Sensor Network and Its Application in Agriculture

Narendran M.[1*] and Prakasam P.[2]
1. Research scholar, Anna University, Chennai, INDIA
2. Director/ CLT, SNS College of Engineering, Coimbatore, INDIA
*narendran.m.2016@gmail.com

Abstract
The current study described the design and implementation of a Wireless Sensor Networks (WSN) for monitoring agricultural environment. The ability of WSNs in detecting as well as observing physical phenomena makes them exceedingly popular recently. The nodes are generally subjected to varying environmental conditions and it is a real challenge in establishing a WSN. Experts should take into consideration energy as well as resources-limited nature of WSNs when building algorithms. For achieving scalability, increased network lifetimes, as well as improvement in energy efficacy, clustering is a significant tool to be followed in WSNs. Considerable work has been undertaken for developing effective clustering protocols, and improvement of the efficacy of clustering is an ongoing issue. The criterion that is utilized for electing cluster heads in certain of the protocols mentioned take into consideration mobility-associated measures for reducing frequency re-election of CHs. Random Competition based Clustering (RCC) scheme mainly focuses at simplicity and stability. Invasive Weed Optimization (IWO) refers to a population-based protocol; that owes its inspiration to nature, i.e., the natural activity of weeds in colonizing as well as discovering appropriate places for growing as well as reproducing. In this work, Lowest-ID (LID) clustering is optimized using IWO with the available energy and the distance considered from neighborhood.

Keywords: Wireless Sensor Networks (WSN), Random Competition Clustering (RCC), Lowest-ID (LID), Invasive Weed Optimization (IWO).

Introduction

India is the second largest agricultural producer of wheat and rice in the world. Farming is one of the main occupation in India. This is essential for feeding the growing population and for exporting wheat and rice. Though India is ranked second worldwide for farm output, poor infrastructure and unorganized farming techniques lead to high food losses. This is the main objective behind designing monitoring systems in agriculture sector for improving the efficacy and yield of the farming. Systems to collect meteorological data, soil condition, moisture, humidity, pest count is required for planning for the right crop, communicating with agricultural experts and so on. Wireless Sensor Network (WSN) is used in particular areas for sensing the information. In the wireless sensor node base station has unlimited amount of energy but sensor nodes available in the network with limited amount of energy. WSN comprises of several thousands of small, less-powerful sensor nodes that are arbitrarily or manually set in the target location. Every sensor node collects local data, processes them, sending it to remote base stations (BSs) known as sinks. The primary limitation of the sensor node is its restricted as well as non-replaceable power source. Power usage for sensor node is the primary challenge in long-term operation of WSN [1]. The main issue in the wireless sensor network is the reduction of power usage. WSNs are utilized to monitor particular areas, particularly locations which are not very accessible like battle fields, volcanoes, forest fires, jungles and so on. They can measure temperature or humidity apart from several other applications [2, 9]. The replacement or recharge of node batteries is not feasible, but it is power usage of nodes that is the primary factor impacting WSN's performance at the time of communications. A disadvantage is the decrease in WSN's life time. A technique for improving lifetimes is the reduction of the quantity of transmissions through cluster method. However, the selection of CHs, their quantity as well as member nodes are an NP-hard problem.

Clustering is a data analysis tool for grouping similar data. It has been used for storing and representing large amounts of information as data. Cluster analysis can be defined as discovering natural hidden groups of objects. It has been used for assigning the same objects to the same groups, where different objects are in different groups. Clustering has been applied in many fields, like engineering, computer science, economics, life and medical sciences, astronomy, earth science and social science. Clustering denotes the splitting of unlabelled data set into sets of like components. All groups, are called as clusters, and comprise components which are like one another, and are different from those of other clusters.

All clusters have cluster heads (CHs) as well as a certain amount of cluster members. The CH frequently notifies others regarding the member nodes through usage of a distinct transmission channel. Intercluster communication is transmitted through CH. For reducing the overheads of cluster head communication, the quantity of clusters should be decreased to a minimum in the network. Thus, cluster heads are arranged so as to encompass every part of the network, as well as all the nodes. This enhances spatial reusage of intracluster transmissions. Hence, almost all cluster-based methods lead to non-overlapped clusters wherein CH has several net- work interfaces with varying transmission ranges (for example, shorter range for intracluster, as well as longer range for intercluster transmissions).

Lowest-ID protocol considers every node as being designated a distinct ID. Periodically, nodes broadcast lists of nodes they can hear, which include themselves. Nodes that solely hear nodes with IDs greater than themselves are CHs, unless they particularly give up their role. The least ID nodes hear are their CHs. But, the protocols have some disadvantages; CHs may delegate their duties to the subsequent node with minimal ID in their cluster; extremely mobile node with least ID as well as chain effect might lead to extreme reclustering. The strategy is biased toward nodes with least IDs that lead to battery draining of particular nodes, as well as does not attempt to balance load in a uniform manner across every node [3].

Optimization issues are present in almost every domain n science, engineering, as well as technology. In several of these issues, it is required to calculate global optima of multi-variable function. The parameters which determine the function to be optimized may be continuous or discrete, and moreover, they frequently are to fulfil particular restrictions. Global optimization issues are a part of the complexity class of NP-hard issues [4]. These are particularly hard to resolve, conventional descent optimization protocols on the basis of local information are not adequate for their solution. In many optimization issues, it is required to assess a huge space for finding global point; the computation procedure consumes too much time. Few of the huge drawbacks of techniques concerned with optimization issues are handling huge search spaces with higher computation costs.

Most ordinary evolutionary methods, like TS, GA, etc., are slow in converge. In recent years, new methods, such as ACO, PSO, ABC and MICA, were introduced to obtain better solutions and converge more quickly. Invasive Weed Optimization (IWO) is one of these new evolutionary algorithms. IWO was applied for clustering and the scores obtained by this method were either less or equal to the other clustering algorithm's scores [5]. Invasive Weed Optimization (IWO), that owes its inspiration to the activity of weeds that grow in nature, has higher exploration abilities, and is capable of converging to the optimum solution of problems in an efficient manner.

In the current study work an Invasive Weed Optimization (IWO) is proposed to optimize the LID clustering based not only the ID but also with available energy and distance from the neighbourhood. Section 2 details the literatures related to this work. Section 3 includes the methods and technique used for proposed work, section 4 represents the results and discussed the obtained results. Section 5 concludes the work.

Literature Survey

In WSNs, the effect of mobility as well as node failures on information delivery as well as network lifetimes are important. Mahapatro & Khilar [6] presented a mobility aware clustering protocol that functions along with error detection protocol. Nodes identified as defective are not eligible in CH selection, which leads to improvement in information delivery. Energy as well as mobility aware greedy geographical routing algorithm is built that creates a CH backbone for intracluster communications. As the stability of hierarchical infrastructure is important for the determination of network performance, sensor nodes elect themselves are CHs on the basis of remaining energy as well as mobility. A non-CH node selects a CH with greatest link stability. For addressing the hotspot issue, the study sorts nodes into clusters of sizes which are not equal.

In clustering, Cluster Head (CH) is responsible to send data via other CH nodes to the base station. This tends to create situation known as nodes are die earlier. To increase the node aliveness in the network, power usage must be decreased in the network. To decrease the energy consumption of energy proper selection of cluster head is required. Pachori & Suryawanshi [7] devised an algorithm for cluster head selection which is NP hard problem. For the selection of cluster head markov model was used which is used to predict cluster head in future based on the current state. The protocol is checked with respect to the no. of rounds vs remaining energy in the network.

Gavalas et al [8] introduced an effective distributed clustering protocol which utilizes both mobility as well as energy measures for providing stable cluster formation. CH is originally chosen on the basis of time as well as cost-effective least-ID technique. At the time of the clustering maintenance stage although, node ID is reassigned as per node mobility as well as energy state, guaranteeing that those with minimal mobility as well as adequate energy are designated with least IDs, and therefore, are chosen as CH. Proposed protocol further decreases control traffic volumes as broadcast periods are modified as per node mobility patterns: employed non-frequent broadcasting for relative static network topology, as well as improve broadcast frequency for high mobile network configuration. Simulations reveal that power usage is uniformly distributed amongst network nodes and signalling overheads are considerably reduced.

A huge challenge of WSN is the improvement in lifetime. Insufficient power of node is the primary challenge. This problem can be solved through optimization of node energy usage. One such method is clustering; however, optimal clustering of WSN is an NP-hard issue. Shanbehzadeh et al[9] proposed a hybridized protocol on the basis of Genetic Algorithm as well as Particle Swarm Optimization for overcoming the clustering issue through discovery of quantity of clusters, CHs, as well as member nodes. Simulations show that the protocol performs better than LEACH, as well as GA based clustering strategy.

Huang et al[10] proposed a 2-stage IWO method. In the 1st, IWO is concerned with exploration for discovering promising solutions With the acquired solutions in this stage, a novel clustering scheme that is employed in this study is utilized for capturing varying promising solution areas. In the 2nd stage, an altered IWO is used for searching all promising areas with care. On the basis of the outcomes of clustering, the value of key variable is defined through statistical data but not artificial settings. In this manner, variable issue is resolved; balance between explorative as well as exploitative stages are attained. Outcomes of experiments reveal that the suggested method is both effective as well as efficient that is capable of both exploring as well as exploiting the promising areas in search space with efficacy as well as get outcomes that are better than the generic IWO.

Chowdhury et al[11] applied an evolutionary meta-heuristic protocol called IWO for automated splitting of data set with no previous data regarding the quantity of natural groups in it. Fitness function utilized in GA is a cluster validity index. Based on the outcomes of the index, IWO gives the split data set together with the adequate quantity of divisions. The efficiency of the protocol is contrasted with the varying string length GA with point symmetry based distance clustering (VGAPS-clustering), varying string length Genetic K-means protocol (GCUK-clustering) as well as a weighted sum validity function based hybridized niching GA protocol (HNGA- clustering) and is represented for 9 artificial data sets as well as 4 real life data sets.

Clustering refers to a non-supervised learning technique which is utilized for grouping like objects. A popular as well as effective clustering technique is K-means, because it has linear time complexity, as well as easy for implementation. But, it might be forced into local optimum. Hence, several techniques are suggested through hybridization of K-means as well as other techniques. Boobord et al [12] suggested a hybridized method which hybridizes IWO and K-means. The IWO algorithm is a recent population based method to iteratively improve the given population of a solution. In this study, the algorithm is used in the initial stage to generate a good quality solution for the second stage. The solutions generated by the IWO algorithm are used as initial solutions for the K-means algorithm. The proposed hybrid method is evaluated over several real world instances while the outcomes are contrasted with popular clustering methods in the canon. Outcomes reveal that the suggested technique is promising in contrast to other techniques.

Methodology

The proposed IWO technique is compared with RCC and LID. 150 nodes each with initial energy of 0.5J and having temperature and humidity sensors are used.

Random Competition based Clustering (RCC): Nodes identification as well as arbitrary numbers are utilized by RCC [13] for clusters formation based on First Declaration Wins Rule. The rules assign governorship position for sensor nodes which declare themselves initially as cluster head to the other nodes within its range. Another cluster formation technique is broadcasting which may be split into direct as well as multihop. Cluster advertisement messages in the direct broadcasting are transmitted from CHs to other sensors in the chosen region. During the reception of the advertisement message, receiver replies to the cluster head and refrains itself from the reception of other advertisement messages. In multihop broadcasting, CHs use particular transmission ranges for sending advertisement messages to other sensors. The receiver has the capacity to determine if it ought to retransmit the received ad message to the other sensor in its transmission range. The multihop broadcasting decreases power usage as there is a communication limit. The sensor nodes that are not near one another are not required to share messages in a direct manner. But, the multihop broadcasting has a drawback on having more delays in comparison to the direct one. The delays are because the data needs to undergo processing in all nodes along the multihop route that leads to delays in cluster building.

Lowest-ID (LID): A popular one is the Lowest-ID (LID), where every node is designated a unique ID. In a periodic manner, node broadcasts its ID via ‘Hello’ control message, in a period called the ‘Hello period’ (HP). The least-ID node in a neighbourhood is chosen as the cluster head; nodes that are able to hear 2 or more cluster heads are turned into gateways, and the rest are taken as member nodes. The primary benefit of LID is the ease of implementing it. It is furthermore, a rapid clustering technique, because it takes solely 2 HPs for deciding on cluster architecture, as well ensures more stable cluster formation than HD. Contrastingly, HD requires 3 HPs for establishing a clustered infrastructure. But, the primary disadvantage of LID heuristic is the bias toward nodes with lesser IDs; the nodes are more probable to function as cluster heads for longer duration leading to quick battery draining. Additionally, neither LID nor HD protocol consider account mobility measures, that is, extremely mobile nodes have same probability to be chosen as cluster heads, though their movement away from their attached member nodes might result in a ripple reclustering impact [8].

Invasive Weed Optimization (IWO): IWO is a recent numeric stochastic optimization protocol. It was developed by Mehrabian and Lucas [14]. The protocol has a simple process with good exploration and diversity. IWO imitates the activity of weeds in colonizing as well as discovering adequate location for growing as well as reproducing in nature. The optimization process is initialized by randomly generating solutions in the space. IWO refers to a population-based metaheuristic protocol which imitates the colonizing activity of weeds for adapting to an external environment. The fundamental characteristics of weeds are that they grow their population completely or mainly in a specific geographic locale that may be considerably huge or tiny. Initially, a particular quantity of weeds are arbitrarily distributed across the whole space [9]. They gradually grow and implement the steps given below:

(1) Initialization: A certain quantity of weeds are arbitrarily initialized within the possible search region.

(2) Reproduction: The quantity of seeds yielded by a weed is dependent on its fitness and the least as well as greatest fitness of the population, as given by equation (1):

$$n_{weeds} = \frac{f - f_{\min}}{f_{\max} - f_{\min}} (n_{bestpro} - n_{leastpro}) + n_{leastpro} \quad (1)$$

(3) Spatial dispersal: The distance between seeds as well as parent weeds obey normal distribution with 0 mean but differing SD given by equation (2):

$$\sigma_t = \left(\frac{t_{\max} - t}{t_{\max}} \right)^{pow} (\sigma_{\max} - \sigma_{\min}) + \sigma_{\min} \quad (2)$$

Then, the location of the jth seed given by the ith weed may be denoted by equation (3):

$$S_{ij} = W_i + \sigma_t.randn(0,1) \quad (3)$$

(4) Competitive exclusion: Solely the most fit weeds are considered in the colony and steps 2-4 are iterated till maximal population size is arrived at. IWO offers a simplistic as well as clear evolutionary method for optimization. All weeds grow in an independent manner; hence, there are several centres that have strong capacity for global searches.

But, the spatial positions of seeds are determined through equation (3) and SD value is got through formula (2), that has its influence due to other variables. The competitive exclusion operator chooses improved solutions solely on the basis of fitness values that might make most new weeds situate in the same solution area as their parents. Moreover, IWO has various needs of SD values in various search phases, and it influences the protocol's performance greatly. The factors typically ensure that IWO is not able to attain the global optima.

Optimized LID using IWO: Clustering protocol is taken up for capturing various promising solution areas of higher quality solutions. The clustering strategy [16] has its basis in distance as well as density.

Cluster Head Selection using IWO
CH Selection protocol
Step 1: Start.
Step 2: For all members in the cluster
Step 3: Compute the distance utilizing x as well as y parameters with other clusters as well as ID.
Step 4: If ID is the least and it is nearer the maximal quantity of nodes, that is, having greatest connectivity with other nodes in the cluster
Step 5: Iterate step 3 and 4.
Step 6: In the end, the chosen CH is obtained with least ID as well as highest connectivity.
Step 7: End.

The proposed method produces hybrids through the IWO algorithm and CH selection. First, IWO is used to produce a good quality solution (the solution that is near to optimal). Then, the output of the IWO is used as an initial cluster centre for the CH selection method. The method exploits the search capacity of the IWO protocol to overcome the local optimum problem of the CH selection algorithm. More specifically, the task searches for a good approximate initial solution for the CH selection algorithm. It is presented the CH selection algorithm in two stages as follows:

IWO Algorithm:
1. Initialize the solution population.
2. Evaluate the fitness of the population.
3. Every member of the population produces seeds according to its fitness.
4. The seeds are distributed across the search area randomly by normal distribution and adaptive standard deviation.
5. The procedure is continued till the maximal quantity of plants is arrived at. Then, solely the fittest plants are capable of surviving as well as reproducing seed, till the maximal quantity of iterations are arrived at. (Refer Step a)

Clustering Algorithm:
6. The best solution obtained from the last stage is selected as an initial solution of the CH as in step a.
7. Assigning objects to clusters.
8. Calculate the new position of K cluster centres.
Repeat steps 7 and 8 until the maximum iteration is reached.

Results and Discussion

LID and Optimized LID schemes are evaluated. Table 1 to 3 and figure 1 to 3 shows the average end to end delays, average PDR, and network lifetime respectively

Table 1
Average End to End delay (sec) for Optimized Scheme

Node mobility in m/s	Lowest ID	Optimized Lowest ID scheme
5	0.0016	0.00151
10	0.00212	0.00155
15	0.01998	0.01507
20	0.02462	0.01889
25	0.05167	0.04688
30	0.05963	0.04043

From the table 1 and figure 2, it can be observed that the proposed techniques has lower average end to end delays by 31.06% for 10 nodes, by 26.34% for 20 number of nodes as well as by 38.38% for 30 number of nodes when compared with LID.

From the table 2 and figure 3, it can be observed that the proposed techniques has higher average packet delivery ratio by 18.89% for 10 number of nodes, by 18.99% for 20 number of nodes and by 29.15% for 30 number of nodes when compared with LID.

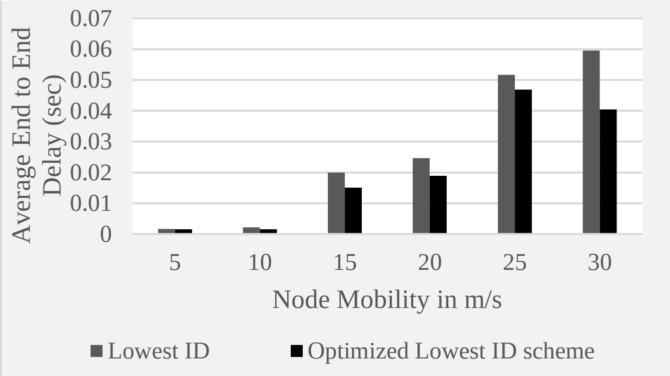

Figure 1: Average End to End delay (sec) for Optimized Scheme

Table 2
Average Packet Delivery Ratio for Optimized Scheme

Number of nodes	Lowest ID	Optimized Lowest ID scheme
5	0.8378	0.9843
10	0.8098	0.9788
15	0.7927	0.9709
20	0.7565	0.9153
25	0.7049	0.857
30	0.6133	0.8226

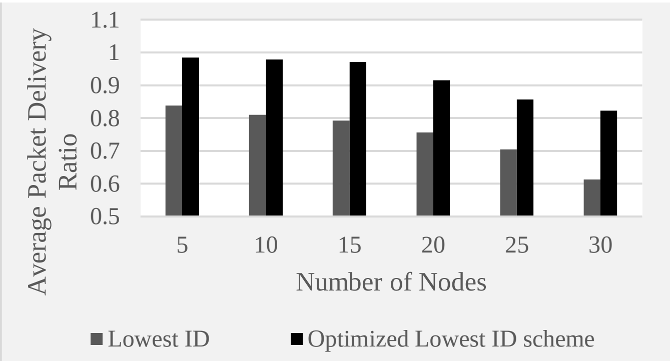

Figure 2: Average Packet Delivery Ratio for Optimized Scheme

Table 3
Network Lifetime for Optimized Scheme

Number of rounds	Lowest ID	Optimized Lowest ID scheme
0	100	100
100	92	96
200	78	90
300	58	87
400	26	81
500	3	65
600	0	44
700	0	23
800	0	11
900	0	2

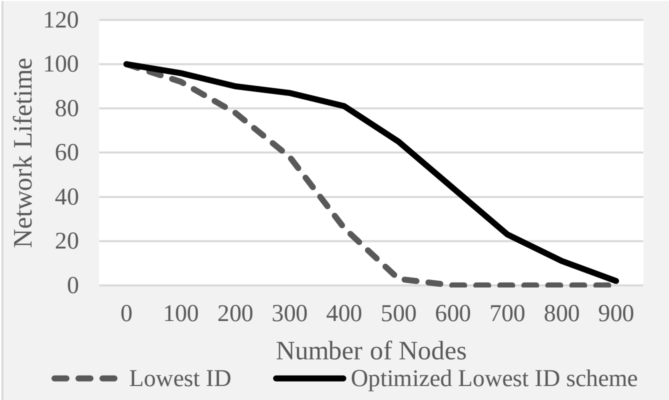

Figure 3: Network Lifetime for Optimized Scheme

From the table 3 and figure 4, it can be observed that the proposed techniques has higher network life time by 4.26% for 100 number of rounds, by 14.29% for 200 number of rounds, by 40% for 300 number of rounds and by 102.8% for 400 number of rounds when compared with LID.

Conclusion

WSNs have obtained great interest for their vast variety of applications. The current study proposes an energy aware WSN for agricultural application. Clustering refers to a promising method to build hierarchies as well as to simplify the routing procedure in mobile adhoc network environments. IWO is an adequate protocol that is on par with other evolutionary protocols. It is easy for understanding as well as programming; has excellent resilience as well as rapid global search capacity. Results show that the proposed techniques has lower average end to end delays by 31.06% for 10 nodes, by 26.34% for 20 number of nodes and by 38.38% for 30 number of nodes when compared with LID. Also the proposed techniques has higher average packet delivery ratio by 18.89% for 10 number of nodes, by 18.99% for 20 number of nodes and by 29.15% for 30 number of nodes when compared with LID.

References

1. Singh A., Clustering Approach in Wireless Sensor Networks, A Review **(2016)**

2. Gupta S.K. et al, Clustering protocols in Wireless Sensor Networks, A Survey, *International Journal of Applied Information Systems*, **5(2) (2013)**

3. Dubey G., Shastri A. and Khandelwal M., Implementation of an Energy Efficient Routing Protocol for WSNs, *International Journal of Computer Applications*, **86(12)**, 13-16 **(2014)**

4. Vali M., New Optimization Approach Using Clustering-Based Parallel Genetic Algorithm, *Neural and Evolutionary Computing* **(2013)**

5. Chowdhury A., Bose S. and Dos S., Automatic Clustering based on Invasive Weed Optimization Algorithm, In Proceedings of the 2nd International Conference, Swarm, Evolutionary and Memetic Computing, Andhra Pradesh, India, 105-112 **(2011)**

6. Mahapatro A. and Khilar P.M., Mobility adaptive unequal cluster–based routing protocol in wireless sensor

networks, *International Journal of Sensor Networks*, **14(2)**, 65-81 **(2013)**

7. Pachori N. and Suryawanshi V., Cluster Head Selection Prediction in Wireless Sensor Networks, *International Journal of Computer Science and Information Technologies, IJCSIT*, **6(2)**, 1033-1035 **(2015)**

8. Gavalas D., Pantziou G., Konstantopoulos C. and Mamalis B., Lowest-ID with adaptive ID reassignment: a novel mobile ad-hoc networks clustering algorithm, *Distributed, Parallel, and Cluster Computing* **(2011)**

9. Shanbehzadeh J., Mehrjoo S. and Sarrafzadeh A., An intelligent energy efficient clustering in wireless sensor networks, In Proceedings of the International Multiconference of Engineers and computer Scientists, **1**, 614-618 **(2011)**

10. Huang S., Ren Z., Wang M. and Sun C., A Two Stages Invasive Weed Optimization via a New Clustering Strategy, In Proceedings of the 2016 on Genetic and Evolutionary Computation Conference Companion, ACM, 1001-1005 **(2016)**

11. Chowdhury A., Bose S. and Das S., Automatic clustering based on invasive weed optimization algorithm, In International Conference on Swarm, Evolutionary and Memetic Computing, Springer Berlin Heidelberg, 105-112 **(2011)**

12. Boobord F., Othman Z. and Bakar A.A., A WK-Means approach for clustering, *Int. Arab J. Inf. Technol*, **12**, 489-493 **(2015)**

13. Mardini W., Yassein M.B., Khamayseh Y. and Ghaleb B.A., Rotated Hybrid, Energy-Efficient and Distributed (R-HEED) Clustering Protocol in WSN, *Wseas transactions on communications*, **13**, 275-290 **(2014)**

14. Mehrabian A. and Lucasc C., A Novel Numerical Optimization Algorithm Inspired from Weed Colonization, *Ecological Informatics*, **1(4)**, 355-366 **(2006)**

15. Pourjafari E. and Mojallali H., Solving nonlinear equations systems with a new approach based on invasive weed optimization algorithm and clustering, *Electrical Engineering Department*, **4**, 33-43 **(2012)**

16. Rodriguez A. and Laio A., Machine learning. Clustering by fast search and find of density peaks, *Science*, **344**, 1492-1496 **(2014)**.