Dominating Set & Clustering Based Network Coverage for Huge Wireless Sensor Networks

Mohammad Mehrani, Ali Shaeidi, Mohammad Hasannejad, and Amir Afsheh

Abstract—**Routing is one of the most important issues in wireless sensor networks. In order to transfer packets in the network effectively, a sound path between nodes should exist to reduce data transmission costs. A proper coverage algorithm is very critical to establish a suitable routing algorithm. Small set of nodes in a network is an important consideration for a coverage algorithm. This paper proposes a problem entitled "inter and intra cluster dominating set based routing". Clusters are made in large networks with numerous nodes to perform coverage in both inter and intra cluster levels. Cluster head as the manager of the cluster divides the cluster area into some cells and selects a node for each cell to work as the router of corresponding cell. Regardless to cluster head, the pivot nodes are also in charge of routing in the network. As the result of this algorithm, all the cluster heads and pivot nodes together make a group called *dominating set* which covers all over the entire region. This is designed to achieve full coverage and also its nodes establish routing backbone of the network. This is carried out in such a manner that every one of the network nodes has direct access to one of the selected pivots. The presented coverage algorithm is based on the platform of our last work in clustering (FEED). Using this technique reduces existing load of cluster head as well especially in huge networks and considerably decreases energy consumption of the regular nodes.**

Index Terms—**coverage, clustering, lifetime**

I. Introduction

To introduce wireless sensor networks some elements are to be mentioned. The first element is that WSNs usually consist of a big number of sensor nodes. A typical sensor node has a processor, a wireless radio and a sensor to sense. Furthermore, a battery is something critical for a sensor node. All the sensors send data that they sense from the given region to base station. They can do it either directly or via some other nodes. There are some very important issues in this action like considering energy efficiency which plays an affective role in wireless sensor networks because these nodes are equipped by batteries with limitations in saving and consuming energy. There are other factors related to energy efficiency directly or indirectly, like having a good routing protocol.

The path selected by nodes to send data is very affective because energy resources are limited and routing protocol should pay attention to this fact and has to try to send data in the path which doesn't lead the nodes to loss much energy. There are lots of routing protocols capable to route the network soundly. Lots of algorithms have been presented to achieve to optimum routing protocols. Having a very fast and energy efficient routing protocol is the basic purpose.

To create an optimum routing protocol there should be some nodes to cover all the network nodes. So, coverage is another important factor in wireless sensor networks. Coverage in this case means the ability of seeing every network node by at least one node of coverage dominating set. To cover whole the network one should select the coverage nodes among all the network nodes in a fast way.

Selecting routing nodes is so important, because they should be able to make relation with all the network nodes even with those of nodes existing far from other nodes. By doing this we cover whole the network. It must be said that router node selection should make minimum amount of network overhead, because overhead problem is also an important and affective factor in WSN and may make some big problems if not considered as well.

The size of wireless sensor networks can be very big such that the number of sensors can be even tens of millions. The words must be said are some coverage algorithms can cover just small networks and make network overhead and broadcast redundancy when the network has so much nodes and the size of the network is big. So, coverage algorithm has to be able to cover whole the network completely and without making network overhead whether the network is big or small. Using clustering techniques helps to cover crowded big networks.

Clustering is a technique used to reduce energy consumption of network nodes. In clustering a node with best situation is to be selected as cluster head. Cluster head selection is different in different kinds of clustering methods. Some of clustering algorithms try to select cluster heads according to a probability formula [1] while some more are eager in energy and distance factors to do this [2]. Another clustering method [3] pays attention to some more factors rather than energy and distance, for example centrality, etc. After choosing cluster head other nodes give data to cluster head and then cluster head forwards data to base station. By doing such, energy consumption of network nodes reduces very well and just cluster head has more energy consumption. So, cluster head is to be selected in a way which has more energy than most of other nodes and also can relate with other nodes easily.

Note that cluster head shouldn't exist in a far part of cluster even if it has much more energy than other network nodes, because in this situation the distances between cluster head and some of the other nodes in the cluster will be so much and will lead these nodes to loss lots of energy if they

Manuscript received February 29, 2012; revised April 20, 2012. This work was supported by computer department of Islamic Azad University, Shoushtar branch, Shoushtar, Iran.

M. Mehrani is a member of scientific board of Islamic Azad University, Shoushtar branch, Shoushtar, Iran. Wireless Networks and\Ad hoc are his research fields.

A. Shaeidi works in the PNU University of Dezfoul as a member of scientific board. He is interested in Ad hoc and Sensor Networks.

M. Hasannejad and Amir Afsheh are both interested in Ad hoc.

want to relate with cluster head. So it'll be against the energy efficiency factor.

In some wireless sensor network implementations when a node losses its ability or when it fails because of different reasons there isn't any node to replace the failed node whether the failed node is a cluster head or a regular node. But some of WSN algorithms are fault tolerant and are able to replace failed or faulty nodes if necessary. If cluster head for example in a given wireless senor network fails then the information of the corresponding region received by cluster members won't be sent to base station meaning that the region full coverage cannot be satisfied. The clustering method used in this work (FEED) is fault tolerant such that a supervisor node replaces the cluster heads when cluster heads fails. It means this algorithm helps well to achieve at full coverage issue.

Selected nodes for covering the network make a set which name is *dominating set.* Our purpose is to have smallest amount of nodes in dominating set such that they make a fault tolerant network. The dominating set can be affective in making a good routing algorithm.

According to the definition presented in [4] the dominating set problem (DS for short) is defined as follows. Suppose there is a weighted graph G = (V, E) presented for a given wireless sensor network such that every node $v \in V$ has contributions with a weight w_v and a positive integer m such that $0 < m \leq |V|$. The dominating set D is a subset containing at most m nodes from V regarding this note that each node $v \in V$ is either itself in D or it has at least one immediate neighbor node and the sum w_v, $v \in D$ is minimized. Selecting the subset D with smallest number of members leads assuming the smallest dominating set.

Outline. In section 2 we mention a distributed coverage method. Then in Section 3 we mention our previous clustering work (FEED) which presents an optimum network clustering. We introduce our Dominating Set & Clustering based Network Coverage for Huge Wireless Sensor Networks in section 4. We finish with concluding remarks in Section 5.

II. Distributed K-Fold Pivot T-Set

The work presented in [4] is named Distributed k-fold pivot t-set algorithm. This algorithm tries to select the smallest group of sensor nodes for covering all over the network. Also the members of this group together are the best nodes to act as routers for routing in the network. Thus the selected nodes are pivot nodes of the network.

At first, algorithm assumes an empty set D which will contain the network pivot nodes, later. Each node selects for example 4 nodes from its nearest neighbor nodes as its subgroup (*t-ball*). Then the node sends a value 1 to all of its neighbors that are in its *t-ball* and a value 0 to other nodes out of the *t-ball.* Thus, for a given *n* node network there would be *n* number of *t-balls.* Every node sends a value *1/n-1* to its*t-ball* members and a value *0* to others. In this round many values are sent and received by all the network nodes. Each node sends *n-1* values and also receives *n-1* values. Then some node *i* adds all the values it has received together and names this total number x_i. At the rest of algorithm each node which *x-value* is equals to *(n-1/n-1) =1* introduces itself as a pivot node and broadcasts a message to other nodes to let them know about this subject.

So, this node adds itself to pivot set *D.* In the next round each node which *x-value* is equals to *(n-2/n-1)* introduces itself as a pivot node and broadcasts a message to other nodes to inform them about this subject. So, this node adds itself to pivot set *D.* The algorithm continues this way again and again. During all the above rounds, some node *i* compares its *t-ball* with pivot set *D* and changes its color to *gray* if it has at least *k* nodes from its *t-ball* that they are in pivot set *D,* too.

This algorithm finishes when all the network nodes are *gray* and all of them have at least *k* nodes from pivot set *D* that they are in their *t-balls*, too. Now the algorithm is finished and all the nodes in the pivot set *D* are selected as routers and they can cover whole the network.

Fig 1 shows an example of executing this algorithm. In this figure all the nodes are gray and pivot set D is:
D= {2, 5, 6, 8, 10}

III. Feed Clustering Algorithm

FEED (Fault Tolerant Energy Efficient Distributed Clustering for WSN) is our last work which aims at clustering wireless sensor network. It has a significant affect in improving the network lifetime. FEED selects the cluster heads by considering some affective and basic factors like energy, density, centrality and distance between nodes.

There are four kinds of nodes in FEED algorithm: CH nodes are cluster heads that work as regular and typical cluster head in the network. PCH nodes are the pivot cluster heads that are also cluster head but have more power and application usages than CH nodes. PCH nodes are the strongest and also the best nodes to work as cluster head. They can be used as main router nodes for routing in WSN. SN nodes are next kind of nodes in FEED that are supervisor nodes for clusters and cluster heads. They are assumed as substitute node for their cluster head when the cluster head fails. Each cluster head has a SN node that replaces the cluster head if necessary. This causes FEED to be a fault tolerant clustering algorithm. Also the presented characteristic of FEED leads this method to be used as the platform of the coverage algorithms. The fourth and last kind of nodes in FEED is regular nodes.

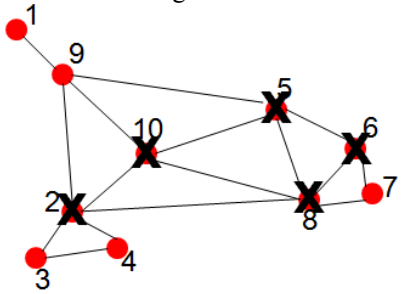


Fig. 1. The result of K-fold algorithm

This algorithm works in four different phases as follows. First phase is designed to calculate the density of nodes and also their centrality. To do such, all the network nodes send their *id* and *coordinate* to all their near neighbours around

and receive same messages from them. Also the angles between neighbour nodes are computed to be used in the centrality counting method presented also by FEED. After the first phase in the condition that not so many messages have been sent and received all the nodes are aware about both of their centrality and density factors. The messages sent in FEED clustering algorithm are usually local broadcast messages without consuming much energy but have direct affect on the FEED performance.

FEED is distributed, so all the network nodes contribute to cluster head election and do it by getting and taking scores from each other. The result of the first phase is calculating the first score made by merging energy, density and centrality factors done by each node. Volunteers in the second phase are those of nodes having enough energy and good density factors. They inform their neighbours about their calculated first scores. It is on all the network nodes to give second scores to volunteers in the third phase. It can be done by entering the distance factors to the received first scores from volunteers. Then they count second scores for volunteers and select the best ones with most second scores as their deputies. Thus, each node sends a message to all of its near neighbours and reveals its deputy volunteer *id* with its calculated second scores.

Now, in fourth phase, each volunteer that has received some second scores from its neighbours calculates its final score by summing the received scores. The volunteer by comparing its final score and the final score of its near neighbour volunteers decides to work as a CH, PCH, SN or even a regular node and informs its neighbours about this decision by sending them a message. These decisions of volunteers are made by paying attention to some notes about the distances between SN, CH and PCH nodes presented in FEED paper. After denoting cluster heads and their supervisor nodes, regular nodes join the nearest ones to make clusters.

According to simulation results, FEED algorithm can prolong network lifetime as well in comparison with LEACH and HEED algorithms that are themselves two well known clustering methods. Also FEED is fault tolerant which is one of the pleasant attributes of this algorithm. Fig 2 shows the network lifetime prolonging made by FEED. It is clear that FEED has significant improvement for network lifetime. Supervisor node replacement can be a reason for this enhancement because in FEED when the energy level of a cluster head falls below a threshold then supervisor node replacement occurs and cluster continues it's functioning until end of the round with new cluster head and without any problem.

Fig 2 shows five hundred percent and also one hundred percent enhancements for network lifetime in comparison with LEACH and HEEDS methods, respectively. Furthermore, the percentage of total remaining energy of the network after 1, 20 and also 50 rounds of three mentioned clustering techniques is revealed by Fig3. This Figure exhibits that when first round finishes, the HEED algorithm has the best performance about network total remaining energy while HEED is replaced with FEED in this Figure during later rounds. All these show that FEED is energy efficient which is a consequent of considering important factors like energy, density, centrality and distance between nodes in choosing cluster heads by this method. Also, cluster head replacement in emergency times is another reason of this successfulness.

IV. Presented Algorithm

In this section we present our network coverage method. As we mentioned before, first of all FEED algorithm performs clustering on the network. After making clusters, cluster head knows about its cluster members meaning their coordinates, energy, etc. Cluster head manages all the cluster operations, so clustering based coverage is managed by cluster head. To achieve at this purpose, cluster head elects the best nodes in different parts of cluster as responsibles of receiving data from the nodes existing in the corresponding part and then to forward it to the cluster head. The elected node is called *pivot node* or *pivot* (for short).

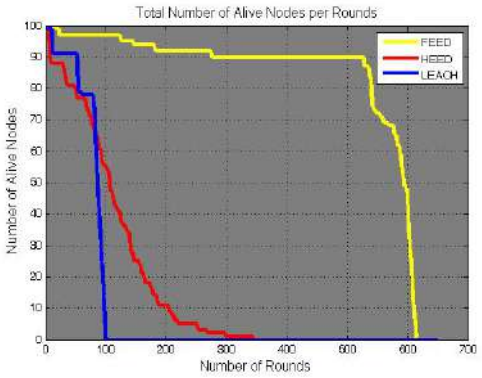


Fig. 2. Total number of alive nodes per rounds

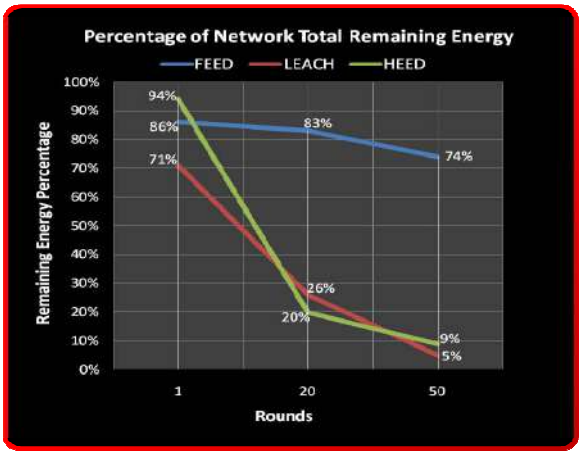


Fig. 3. Percentage of total remained energy of the network nodes

Each pivot should have the best condition in its part meaning that pivot's remaining energy and centrality factors should be the best in comparison with other nodes of the part. Cluster head first assumes its corresponding cluster region as a grid shape. There are some different grid shapes with different affects on the proposed coverage functionality that will be mentioned at the rest of the paper. Then, for each cell of the grid, cluster head elects the best node as the pivot. Cluster head introduces the elected pivots to all the cluster members and also neighbor cluster heads. By doing such, a local map of routers will be made. Also, cluster head assumes two spare nodes for each pivot. Theses spares are second and third best nodes after pivot in their part. This

increases the system reliability and leads to a fault tolerant coverage method because the mentioned spare can be replaced to the pivot in emergency situations like when pivot has been damaged or lost its energy.

Now, the chosen pivots in all of the clusters are routers for routing in all over the wireless sensor network and also for coverage the network nodes. In each cluster, both of cluster head and its supervisor node are in charge of sending received date from the cluster members to the base station. Meanwhile, they aggregate the initial received data and then send them to the base station.

As mentioned before, this algorithm has more applications in crowded networks. In such networks, a huge amount of data is transmitted from regular nodes to cluster head, so the load on the cluster head increases rapidly and can make some problems for cluster head such as losing its energy and correct functionality. By using the proposed method, regular nodes first send data to pivots and then the pivots send data to the cluster head after aggregation.

An example of this algorithm guides to better understanding. Fig 4 shows a part of a crowded and huge network clustered by FEED algorithm. In this figure regular nodes are shown by black stars while cluster heads are exhibited by blue circles and their supervisor nodes are displayed by red circles.

We first focus on a cluster to show the functionality of our coverage algorithm and then pay attention to all over the entire network. Fig 5 reveals a cluster of the network clustered by FEED with its cluster head (blue node) and supervisor node (red node). We explain the algorithm with just one cluster for more clearness.

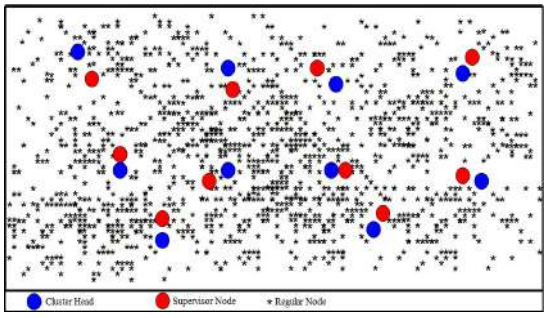


Fig. 4. A part of a huge network clustered by feed method

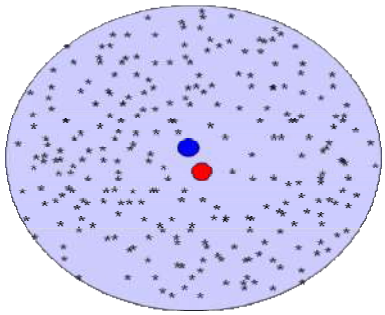


Fig. 5. A cluster with its cluster head and supervisor node

Suppose a cluster of the network that its dominating set is to be made by presented new coverage method. This method is fault tolerant meaning that whenever an important active node fails a replacement will be done to remove the bad affect of this failure. This can be done by getting help from spares. In a large crowded network with big clusters there should be more than one node in charge of covering and routing. It means the best and most powerful nodes should do the cluster head some favors to increase covering performance. As mentioned before, cluster head knows about the position and energy of its members. As shown in Fig 7 By applying this algorithm cluster head assumes its corresponding area as a grid to elect a pivot for each cell of the grid.

Fig 7 exhibits the different states of the grid that can be assumed for the cluster. According to these figure's states, the cluster region is divided into 5, 9, 17, 25, 33 and 45 distinguished parts each one having a pivot with two spares that can replace the pivot if necessary. In these figure, pivots are shown by blue circles while the first and second spares are appeared by red and yellow circles, respectively. Note that cluster head is assumed also as a pivot and its supervisor node as the first spare for central part of the grid.

By doing such, we can make sure that load on the cluster head will be divided on the pivots very well. The percentage of load reduction on the cluster head is related to the number of nodes of the cluster. It is clear that whatever the number of nodes existing in a cluster increases then the load of cluster head increases, too, because the regular nodes send the region data to the cluster head. But, using the grid network divides the load of the cluster head on the pivots.

Fig 6 shows the reduction of load on the cluster head after using different kinds of grids with different number of separate parts and different number of nodes existing in the cluster. According to this figure, the cluster head's load decreases between 10% and 97.5% which is a very significant outcome.

Fig 6 shows that the smaller number of assumed parts is more affective in cluster head's load reduction because whatever this number is bigger the number of pivots is bigger which means the number of nodes that send data to cluster head is bigger which leads having more load on the cluster head.

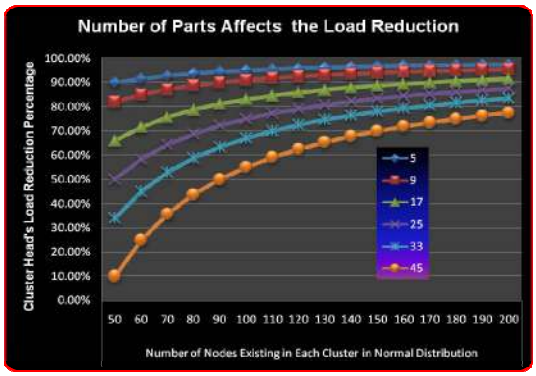


Fig. 6. Load reduction on the cluster head

But, bigger number of pivots leads saving more energy by regular nodes to send data to pivots, because in this condition pivots are more accessible and nearer to contact. It is a consequent of having smaller parts.

Using each one of the different presented grids by Fig 7 will have a different affect on functionality of proposed coverage algorithm. This functionality is related to the distance between part's pivot and cluster head. The nodes existing in the cluster head's part won't have any energy saving because they send data directly to cluster head. Those of nodes who are near to cluster head in the first level parts

have a few energy saving but it is roughly good in comparison with usual situation. The nodes belonging to the second level and later levels would have a satisfying energy saving. Energy consumption for sending a message has direct relation with square of distance that the message takes. Thus, using multi partitioned form for a clustered area will lead some significant affects in energy saving for network nodes.

Cluster members send their received region data to the part's pivot with at most a side of the squared part distance instead sending it to cluster head which distance is much bigger than the length of mentioned side. Note that this strategy for energy saving can be achieved just by using FEED clustering algorithm because this method is distributed such that cluster heads are aware about their member information very well.

Using the different mentioned types of grid shown in Fig 7 leads satisfying results as follows. Applying Fig 7 *a, b, c, d, e* and *f* have 35%, 78.22%, 72.9%, 60.72%, 73.87% and 93.56% energy saving, respectively. Also these results can be seen in Fig 8.

Based on Fig 8 the minimum percentage of energy saving is when the there are 5 separate parts in each cluster. In this circumstances there is about 35% energy saving for almost all the normal nodes but pivots. In other hand, the maximum possible amount of energy saving occurs when a grid with 45 separate parts is used. This leads more than 93% energy saving for the network which is a very satisfying outcome.

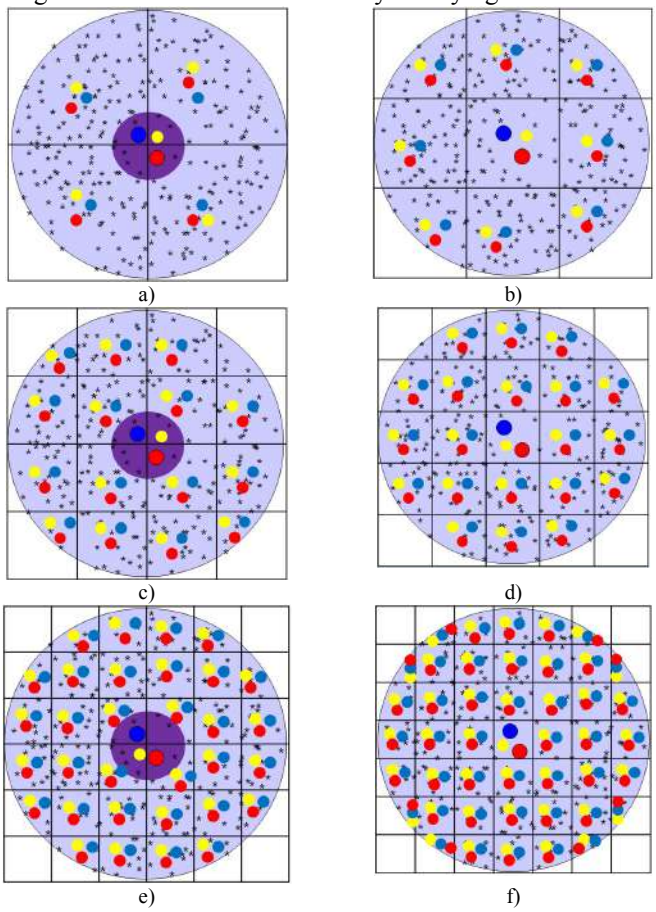


Fig. 7. A partitioned cluster with different grids

In Fig 7 each part has a pivot and two spares. To calculate the reliability of system, suppose that each node fails with a probability of p. So, the reliability of each part is:

$$\text{Re}liability = 1 - p^3$$

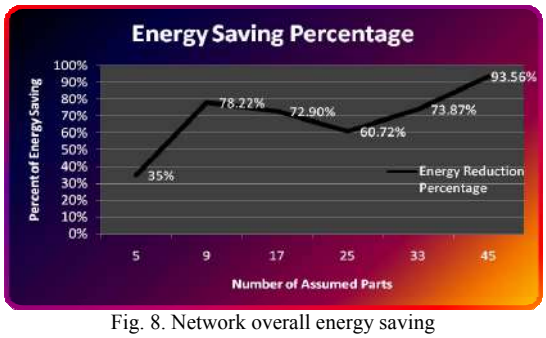


Fig. 8. Network overall energy saving

This is a pleasant reliability for each part and helps the network to not to lose its functionality after possible failures. According to the calculated reliability, in each part the probability that the pivot, first and second spare fail is very few. So, the probability that the received data of a part cannot be sent to cluster head is very few meaning that full coverage is guaranty.

After using the proposed coverage method the network shape is like Fig 9. This figure shows four kinds of nodes: pivots that are shown by blue circles, first spares exhibited by red circles, second spare that can be seen by yellow circles and finally regular nodes displayed by black stars. It is clear that there is a pivot with two spares for each given part of the region. So, achieving at full coverage is guaranty meaning that all the received region information can be sent to the base station. Also the dominating set (pivots) is to be used as a fault tolerant routing backbone. Inter and intra cluster routing is satisfied easily by this algorithm.

V. Conclusion

For having a suitable routing algorithm a proper coverage algorithm is very critical. Some important notes should be considered to make an optimum coverage algorithm: having a small set of nodes that are selected from all the nodes in a network to be accessible for all the network nodes. By doing so, reduction of energy consumption in nodes is achieved. This paper proposed a network coverage method based on dominating set and clustering. Gridding the network and considering a pivot with two spares for each part of the grid is managed by cluster head. Using this method leads very good outcome: reducing load on the cluster head between 10% and 97.5% and saving energy between 10% and 93% for regular nodes are the major consideration and design characteristics which made it as a sound system to be applied in industry.

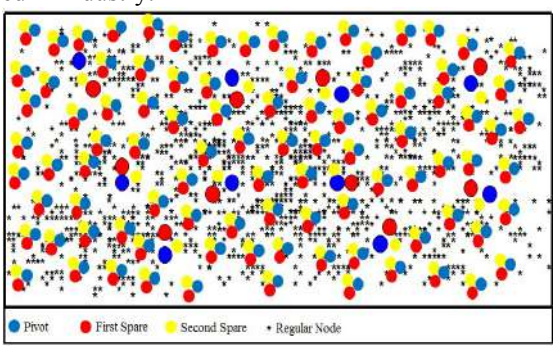


Fig. 8. The region afetr applying new coverage algorithm

REFERENCES

[1] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proc. of the 33rd Annual Hawaii International Conference on System Sciences (HICSS)*, pp. 3005 – 3014, 2000.

[2] O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, DistributedClustering Approach for Ad-hoc Sensor Networks," sponsored in part by NSF grant ANI-0238294 (CAREER) and the Schlumberger Foundation.

[3] M. Mehrani, J. Shanbehzadeh, A. Sarrafzadeh, S. J. Mirabedini, and C. Manford, "FEED – Fault tolerant Energy Efficient Distributed Clustering for Wireless Sensor Networks," *The 12th International Conference on Advanced Communication Technology*, Phoenix Park, Korea, 2010.

[4] D. Bein, "Fault-tolerant k-fold Pivot Routing in Wireless Sensor Networks," in *Proc. of the 41st Hawaii International Conference on System Sciences*, 2008.

[5] K. Matrouk and B. Landfeldt, "RETT-gen: A globally efficient routing protocol for wireless sensor networks by equalising sensor energy and avoiding energy holes," *Ad Hoc Networks* 7, pp. 514–536, 2009.

[6] F. Dai and J. Wu. "Constructing k-connected k-dominating set in wireless networks," in *Proc. of IEEE International Parallel and Distributed Processing Symposium* (IPDPS), 2005.

[7] F. Dai and J. Wu. "Constructing k-connected k-dominating set in wireless ad hoc and sensor networks," J*ournal of Parallel and Distributed Computing*, pp. 947–958, 2006.

[8] C. Glaber, S. Reith, and H. Vollmer, "The complexity of base station positioning in cellular networks," *Discrete Applied Mathematics*, pp. 1–12, 2005.

[9] F. Kuhn, T. Moscibroda, and R. Wattendorf, "Fault-tolerant clustering in ad hoc and sensor networks," in *Proc. of the 26th IEEE International Conference on Distributed Computing Systems*, pp. 68, 2006.

[10] L. D. Penso and V. C. Barbosa, "A distributed algorithm to find k-dominating sets," Discrete Applied Mathematics, pp. 243–253, 2004.

[11] K. Akkaya and M. Younis "A survey on routing protocols for wireless sensor networks" *Ad Hoc Networks*, pp. 325–3, 2005.