2-MASCLE - A Coverage Aware Clustering Algorithm with Self Healing Abilities

Jakob Salzmann, Ralf Behnke, Dominik Lieckfeldt, Dirk Timmermann

University of Rostock
Institute of Applied Microelectronics and Computer Engineering
18051 Rostock, Germany
{jakob.salzmann, ralf.behnke, dominik.lieckfeldt, dirk.timmermann}@uni-rostock.de

***Abstract*— This paper investigates organization problems of large wireless sensor networks. In spite of their random deployment, nodes have to organize themselves as energy efficient as possible to avoid redundant sensor and transceiver tasks. In addition to energy awareness, the network has to guarantee complete sensor coverage and connectivity as long as possible.**

This paper presents a novel clustering algorithm which combines the advantages of horizontal and vertical network fragmentation by introducing a network division in a dual phase cell system. The assisting ability of adjacent cells is exploited to switch-off half of the network cells and allows implementing a self healing algorithm. We compared our developed algorithm to former covering and clustering algorithms and achieved an increased network lifetime compared with them of approximately 80%.

I. Introduction

The recent development of small wireless communication and computing systems makes it feasible to construct sensor network systems for environment surveillance. Such networks consist of hundreds of tiny sensor nodes, each equipped with sensor and wireless communication system, battery and microcontroller.

Sensor nodes are deployed randomly over an area to observe it. In scenarios like forest fire surveillance, precision farming or the measurement of volcanic activity, a network of simple sensor nodes is much more powerful than some single complex sensor systems in terms of detecting phenomena.

However, sensor nodes have to act self-sustained and independently from the environment they observe. Hence, they only have limited energy resources for detection, computation and transmission and 'die' after their resources are exhausted. In order to achieve a maximum network lifetime, the whole network has to act as energy-efficient as possible.

Current researches try to increase network lifetime by different approaches, like energy efficient routing [7] or load balanced communication [8].

Manuscript received July 31, 2007. This work was supported in part by the German Science Foundation (Keyword: Organic Computing).

In this paper, we develop a Clustering algorithm based on the XGAF (Extended Geographical Adaptive Fidelity) algorithm [1]. With a new calculated grid size, we are now able to reduce the number of active nodes in the network. Additionally, we are able to implement an assistance algorithm for neighboring cells, which guarantees a further lifetime extension. The resulting algorithm is called 2-MASCLE (Dual phase – Mutual Assistance in a Cluster Environment)

The remainder of the paper is organized as follows: Section II describes the related work and the already existing XGAF algorithm, Section III presents our new clustering approach, Section IV introduces the abilities for neighborhood assistance, Section V compares characteristic parameters of 2-MASCLE and former algorithms by simulations. Section VI gives a conclusion and an outlook.

II. Related Work

The central problem in wireless sensor networks is power consumption. A major research goal is to increase the total lifetime of a network by temporary deactivating a subset of nodes. Today, there are two different strategies which allow the deactivation of nodes which are not mandatory for a working network with full coverage.

A. Clustering

There exist clustering mechanisms which aspire to a horizontal fragmentation into groups of nodes, which are closely located to each other and cover the same area. Within these groups a clusterhead will be elected, which takes the main functions and may cause the remaining cluster-nodes to sleep. The amount of energy conservation depends on the intra cluster rotation strategy, which is not part of this work. An established cluster-algorithm is LEACH (Low Energy Adaptive Clustering Hierarchy) [4], which first selects clusterheads in a random manner and groups the remaining nodes to these clusterheads.

LEACH, as one example of common clustering algorithms, does not take location into account and creates orderless non-uniform clusters. Therefore it does not guarantee full network coverage. An approach which is more up-to-date and may ensure a full coverage is XGAF which is described in [1].

XGAF clusters the WSN (Wireless Sensor Network) into a regular grid, based on the position of each node. Not only the transmission range but also the sensing ranges of the nodes are used for calculating the geographical cluster size. In contrast to other clustering approaches like GAF (Geographical Adaptive Fidelity) [5] and (E)HGAF (Enhanced Hierarchical Geographical Adaptive Fidelity) [6], it is guaranteed that each member of a cluster is able to sense the whole cluster and not only serves as connector between neighboring clusters.

B. *Coverage*

In contrast to common clustering algorithms, coverage strategies make a temporal apportionment of the nodes. We propose to call it vertical fragmentation. The aim of cover-algorithms is to split the nodes into a various number of groups, called covers, like clustering does. As opposed to clustering, each group should contain nodes from all over the sensor-field. So each group should be able to observe the entire environment. In addition, each cover should be capable of transmitting information from each point of the network to the sink and also to every other node, depending on the used routing-strategy and the networks purpose. Divided into such covers, only one cover has to be online at any time. This facilitates an economy of (n-1)/n of the networks energy, where n is the number of covers. In this case, the WSN would work in n phases.

A random-based simple cover-algorithm is introduced in [2]. In this algorithm each node decides independently which cover it belongs to, only by calculating a random number, which indicates a specific cover. The only requirements for this algorithm to work are an existing random number generator, synchronized clocks and knowledge about the number of covers. The described algorithm is quite simple and needs no additional communication but it has one penalty. The algorithm is not able to guarantee that at least one group covers the entire network. Therefore, some parts of the environment may be left unobserved for a long period. Nevertheless, each observable part of the field will have been observed at least one time after a full cycle of all covers. It is shown in [2, 3] that the algorithm in general performs well and is also inured to clocks which are not exactly synchronized. To achieve a uniform dispersal of the nodes onto all covers and to get covers, which are capable to observe as much as possible of the environment, several algorithms are proposed in [2, 3]. But these algorithms have a high amount of communication.

In addition, regardless which of these algorithms is considered, they all have one fundamental problem. If the nodes are non-uniformly distributed there arise thin populated regions. This phenomenon forces the network to either build only fewer numbers of covers, which are all capable of covering the entire environment, or to build a fixed number of covers assenting that some regions are only part of a subset of covers or some nodes are members of more than one cover.

Furthermore, all of these algorithms do only base upon the transmission range and do not take the sensing range into account. In the given references it is assumed that sensing range and transmission range are equal.

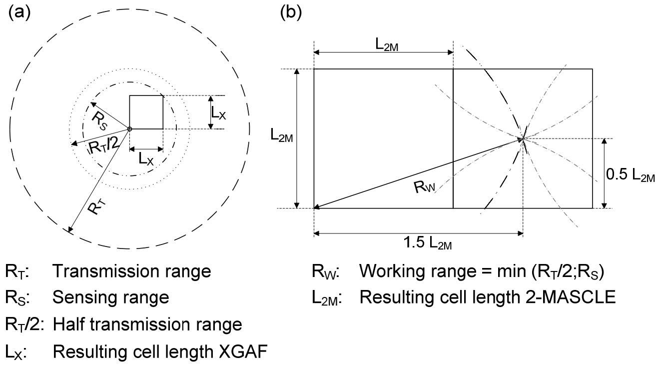

Fig. 1. Calculating the resulting cell length of a) XGAF cell. Resulting cell length guarantees connectivity to adjacent and diagonal adjacent nodes and complete cell coverage by a member node b) 2-MASCLE cell. Resulting cell length allows complete coverage by adjacent nodes and a transmission to one cell above the next adjacent node in a row.

III. Two-phase active Clusters

In this section, we present our approach to combine horizontal and vertical fragmentation which exploits the advantages of clustering and coverage algorithms. Similar to XGAF, we divide the network into clusters. Members of a cluster have to be positioned in the corresponding square cell. Similar to cover-algorithms, we are able to switch-off half of the remaining active nodes. In contradiction to classic cover-algorithms, our algorithm is able to guarantee a complete coverage and connectivity of the network as long as enough nodes are deployed.

To achieve this guarantee, we had to re-calculate the cell size of an XGAF cell, depicted in Fig. 1 a). Instead of reaching the diagonal edge of a cell, the sensing range or half of the transmission range now has to reach the center of an adjacent cell. The new cell length given in (1) is pictured in Fig. 1. b).

$$L_{2M} = \frac{\min(R_T/2; R_S)}{\sqrt{2.5}} \qquad (1)$$

L_{2M} is the new cell length of our algorithm, R_S the sensing range and R_T the transmission range. Just as in XGAF, the cell size allows that only one active node per cell is needed for a complete coverage and connectivity of the network. Additionally, our new cell size allows that the task of an inactive cell can be assigned to its adjacent neighbors. To exploit this ability, the cells are divided into groups.

Based on the calculated cluster size L_{2M} each node autonomously calculates the cell and the corresponding cell group it belongs to by solving a couple of simple equations as illustrated in Fig. 2. As shown in the last line, the cell group either is 0 or 1. This division into two groups according to the presented algorithm leads to a checkerboard pattern of cells belonging to group 0 and those belonging to group 1.

clusterposition_x = position_x div clustersize
clusterposition_y = position_y div clustersize
clustergroup = (clusterposition_x + clusterposition_y) mod 2

Fig. 2. Group selection algorithm using the real node position and the calculated cluster size to identify the cluster and the corresponding cluster group.

The whole development process of 2-MASCLE is shown in Fig. 3. A sensor network with randomly deployed nodes, as apparent in Fig. 3 a), can be divided by horizontal fragmentation like XGAF does, which is depicted in Fig. 3 b) or by vertical fragmentation in Fig. 3 c) like covering algorithms do. With our recalculated cell size, we are now able to combine the advantages of horizontal and vertical fragmentation.

The resulting checkerboard pattern of 2-MASCLE is illustrated in Fig. 3 d). Every second cell is switched off while still a complete connectivity and coverage is warranted.

Using only half of the available cells offers the possibility of running the network in 2 phases like covering algorithms do. After each time T, every cell toggles its active state between active and inactive. This simple algorithm is illustrated in Fig. 4 a).

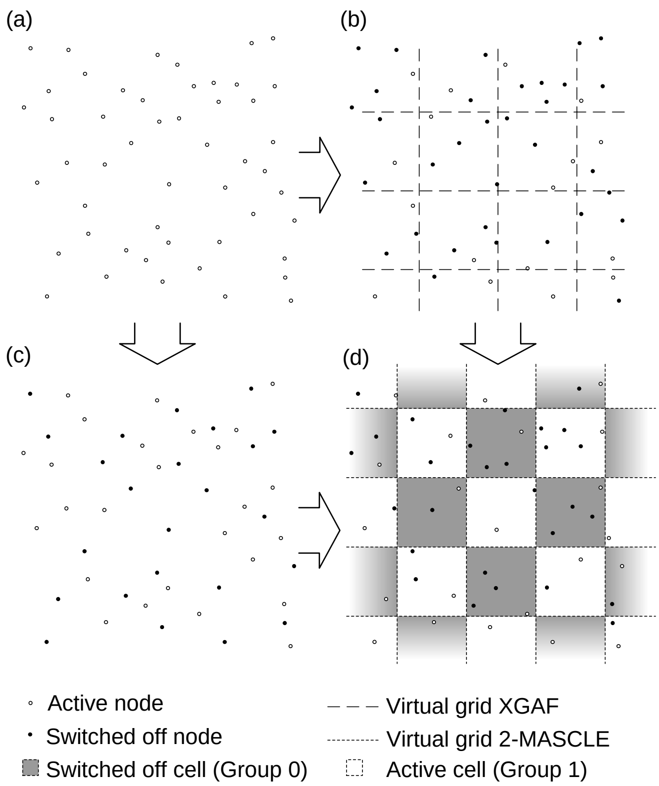

Fig. 3. Development of 2-MASCLE a) Sensor node field with randomly distributed nodes b) Applied XGAF algorithm c) Applied 2-Covering algorithm d) Applied 2-Mascle

IV. Neighborhood Assistance

As described in section III, we have chosen a grid size which facilitates a two-phase scheduling while observing the entire environment during each phase. This effect is based on the fact that the corresponding sensing area of each cluster in the network can be observed by its 4 adjacent clusters. Under this constraint, we are also able to modify the nodes behavior in that way that an empty or dead cell can be compensated by its direct neighbors. To achieve such a behavior, we modified our scheduling algorithm as shown in Fig. 4 b).

After its own active phase, each cell, respectively the clusterhead, stays awake for a certain time T_a and receives wakeup messages from its neighbors. If every neighboring clusterhead signals that its cell is awake the node switches into inactive mode and the cell is switched off. Otherwise the cell is also active in this state. If a cell awakes from inactive state, the clusterhead broadcasts an awake message to its direct neighbors to inform them that its cell is alive.

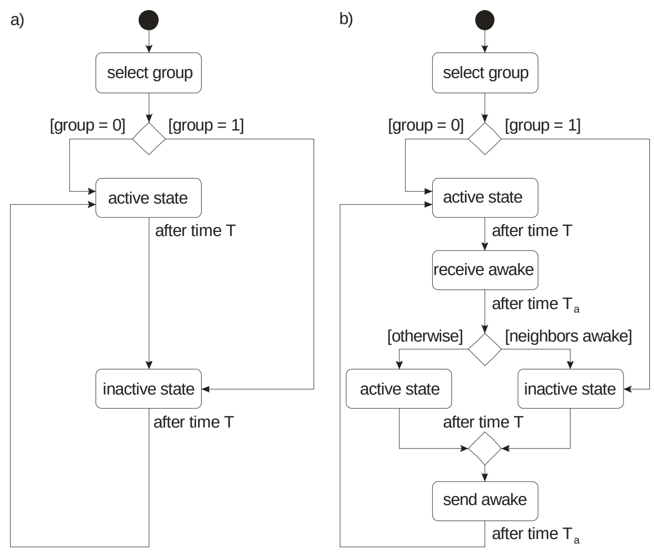

Fig. 4. 2-phase scheduling algorithm for each node a) without neighborhood assistance b) with neighborhood assistance

V. Simulation Results

To demonstrate the strength of our presented algorithm, we have compared our algorithm to those it is based on, i.e. XGAF and cover-algorithm. To maintain statements about lifetime, number of necessary and number of active nodes, we utilized a self contained simulation area, which avoids the influences of edges and corners as shown in Fig. 5.

The first simulation investigated the number of nodes, which are necessary to guarantee the work of the algorithms. Fig. 7 displays the average coverage of a network against the number of randomly deployed nodes.

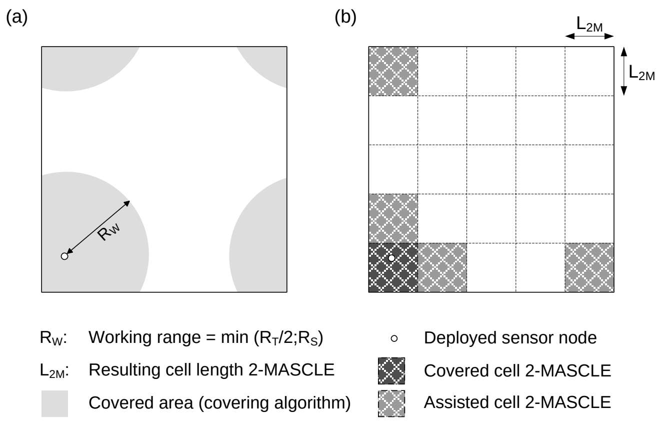

Fig. 5. Simulation area without edge conflicts a) Covering algorithm b) 2-MASCLE

Caused by the fact that a node in 2-MASCLE and XGAF is only switched off if it is nearly unimportant for a full coverage, the performance of both algorithms is near the performance of a not-fragmented network. The 2-MASCLE algorithm without the additional assistance performs similar to the 2-Covering algorithm because in each phase only half of the nodes is available. This is equivalent to the case of 2-Covering.

For comparison, we also displayed the coverage of XGAF and the variations of 2-MASCLE with cell based coverage only. This means an active node can only cover its own cell and parts of adjacent cells which may be assisted by the node's cell but not the whole working range of the node itself.

This simplified view of the network wastes a lot of performance and underestimates the performance of XGAF and 2-MASCLE, so the cover algorithms have to be compared with real coverage of active XGAF or 2-MASCLE nodes. The differences between real coverage and cell based coverage are depictured in Fig. 6.

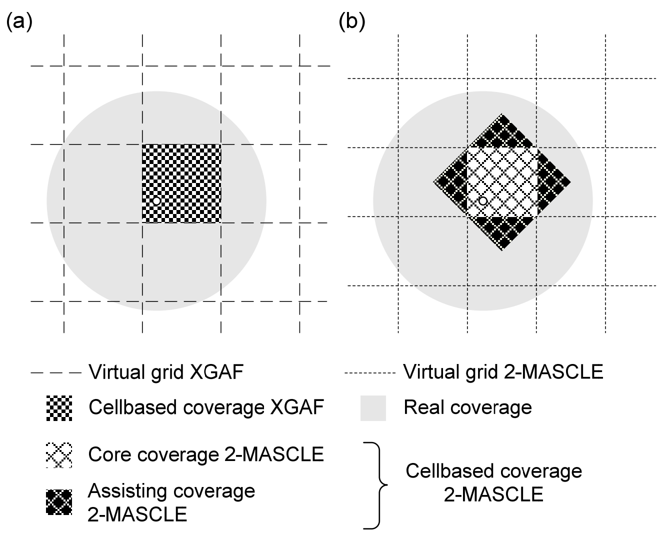

Fig. 6. Real coverage of a node versus simplified cell based coverage of a) an XGAF cell b) a 2-MASCLE cell

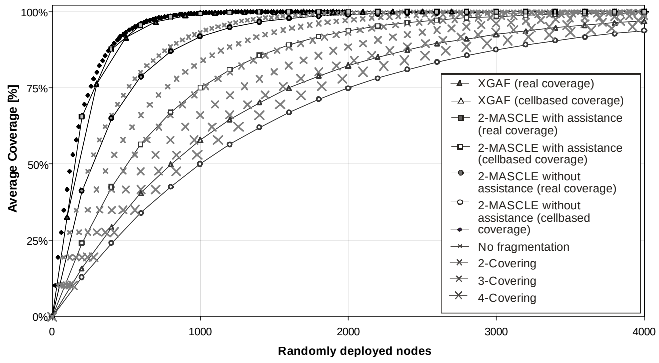

Fig. 7. Average coverage versus randomly deployed nodes. Observed area size: 646m*646m, Working range: 26,87m, Simulations: 500

For a real working network, not only the average coverage of a network but also the average number of networks with complete coverage is important, as shown in Fig. 8.
As expected, 2-MASCLE and XGAF achieve similar results like a network without fragmentation, although 2-MASCLE achieves better results than XGAF. Solely 2-MASCLE without assistance achieves less performance than expected by the results shown in Fig. 7. This is caused by the missing ability of 2-MASCLE without assistance, to cover an empty cell by all eight adjacent cells.

Without an assistance algorithm, only the four diagonal adjacent cells are able to help covering an empty cell, which may often lead to a not fully covered network.

The cell based approaches for calculation of the coverage again need a too high node count. Therefore, they are not considered in further simulations.

The next simulation investigates the number of active nodes versus the area to observe while assuring a full coverage. It can be seen in Fig. 9 that the cover algorithms need at least a 4 times higher number of active nodes than 2-MASCLE, and all algorithms increase linearly with the area. The differences of the active nodes in the *n*-covering algorithms are caused by the fact that each of the *n* phases has to achieve a full coverage, i.e. the overall probability of full coverage is reduced with higher *n*, although n-times more nodes are deployed. Hence, more nodes have to be deployed to achieve the same coverage probability, which results in a higher active node count.

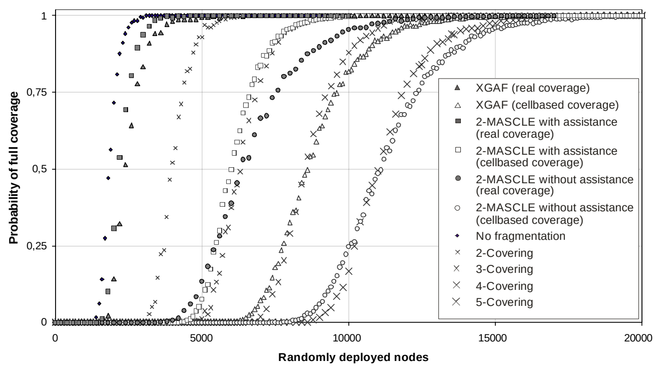

Fig. 8. Probability of full coverage versus randomly deployed nodes. Observed area size: 646m*646m, Working range: 26,87m, Simulations: 500

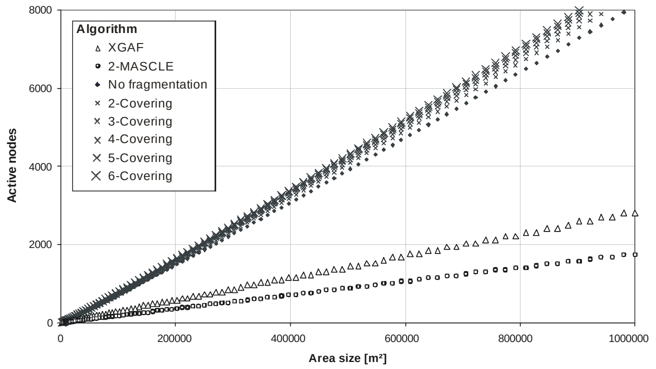

Fig. 9. Active nodes versus size of the observed area. Working range: 26,87m Achieved coverage: 99,9 %, Simulations: 500

The lower number of active nodes is an important fact of 2-MASCLE, which is mirrored in the most important network parameter, the network's lifetime.

Caused by the fact that our algorithm has the goal to maintain a complete coverage as long as possible, a network can be termed as "alive", as long as a full coverage is granted.

Assuming that all active nodes have same energy consumption and switched-off nodes consume a negligible amount of energy, we do not have to consider different node lifetimes and can easily compare the algorithms.

After the battery power of active nodes is exhausted, a new time cycle begins and the network chooses new active nodes via algorithm. The number of active and remaining nodes after each time cycle is presented in Fig. 10. As depicted, XGAF needs a very high number of nodes to grant complete network coverage, so that the nodes are depleted fast.

In comparison to the other algorithms, 2-MASCLE without assistance provides the network with active nodes for the longest time. 2-MASCLE with assistance starts with the same number of active nodes, but after a few time cycles the algorithm tries to heal uncovered areas by activating additional cells. Hence the number of active nodes even increases until no more nodes are available and the network breaks down very fast. The importance of this behavior for the probability of a network, to continue its work with a complete coverage after each time cycle, and hence the lifetime, is presented in Fig. 11.

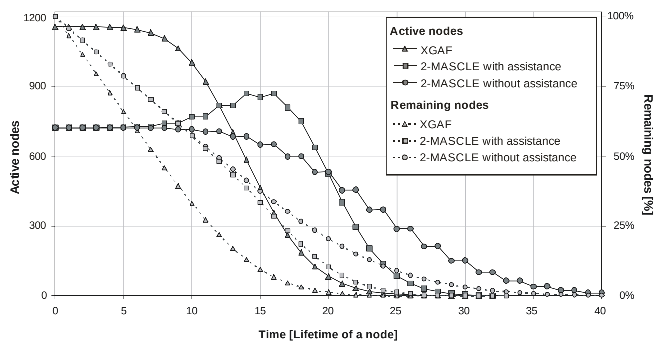

Fig. 10. Active and Remaining nodes versus runtime of the network. Observed area size: 646m*646m, Working range: 26,87m, Simulations: 500, Nodes: 17000

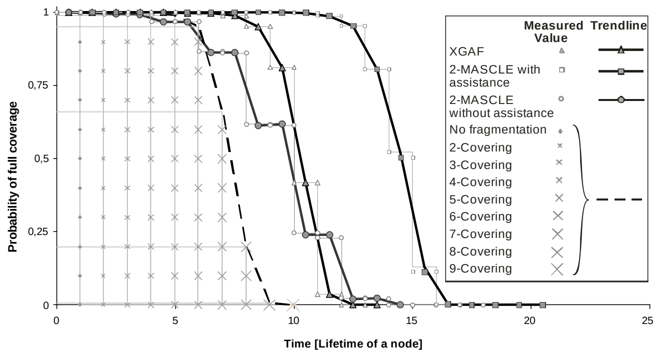

Fig. 11. Average networks with full coverage versus the runtime of the network. Observed area size: 646m*646m, Working range: 26,87m, Simulations: 500, Nodes: 17000

It is obvious, that 2-MASCLE with assistance not only outperforms the covering- and the XGAF algorithms, but also the 2-MASCLE algorithm without assistance in the term of lifetime. Additionally to the before mentioned results, especially in Fig. 11, it is important to know that the covering algorithm is not able to adapt itself to the deployed nodes. Hence one of the covering algorithms has to be selected before deployment. A dynamic selection would be aligned with high communication effort.

To obtain a further quality factor for node management in a WSN, we introduce a term called node efficiency *NE*, which is calculated in (2).

$$NE = \frac{A_c}{N_{act} * \pi * R_W^2} \qquad (2)$$

The term A_C represents the size of the covered area, N_{act} the number of active nodes and R_W the working range of the nodes. The node efficiency denotes the ratio between the actually covered area and the area, which theoretically can be covered by the sum of the node's coverage areas. The maximum node efficiency emerges if the active nodes are distributed in a way that adjacent nodes always form the edges of an equilateral triangle [9]. This optimal distribution of active nodes is also depicted in Fig. 12.

The node efficiency of this scenario is calculated in (3) and reaches a value of about 82.7%.

$$NE_{Max} = \frac{6\left(R_W^2 * \sqrt{3}/4\right)}{\pi * R_W^2} = \frac{3\sqrt{3}}{2\pi} \qquad (3)$$

The node efficiencies of our developed and examined algorithms versus the deployed nodes are depicted in Fig. 13. There it is shown that the *n*-covering algorithms always loose their node efficiency with more deployed nodes and the number of covers *n* has to be continuously adapted to the deployed nodes to avoid too low node efficiency. The diagram also reveals that the 2-MASCLE algorithm with and without assistance aspire to the same value, which is calculated as 25.46%.

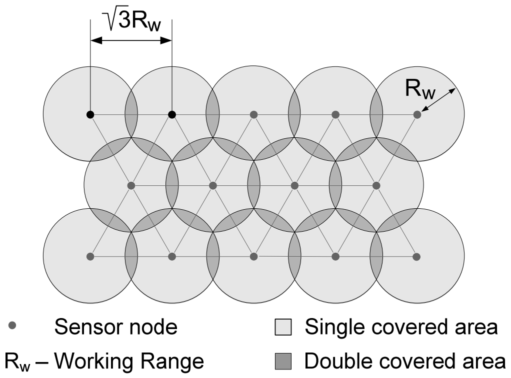

Fig. 12. Deployment of sensor nodes as edges of equilateral triangles for maximum efficiency

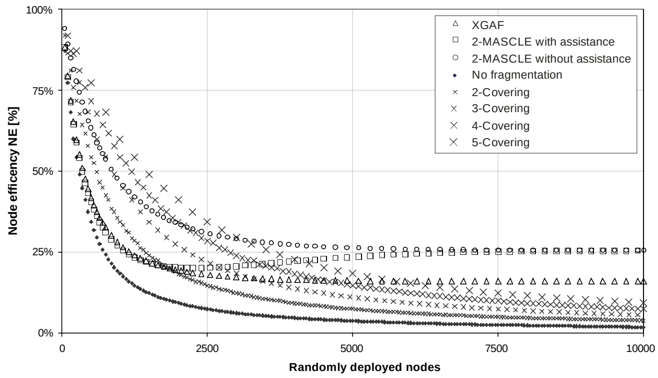

Fig. 13. Node efficiency versus deployed nodes. Observed area size: 646m*646m, Working range: 26,87m, Simulations: 500

This value is fixed if a full coverage is achieved. It complies with the ratio between cell based coverage and real coverage as depicted in Fig. 6. According to the same calculation, XGAF solely achieves a node efficiency of 15.92%, which is also shown in Fig. 13. Comparing XGAF and 2-MASCLE, 2-MASCLE achieves 60% higher node efficiency. As it is shown in the diagram, the node efficiency is closely related to the number of active nodes.

VI. Conclusion and Outlook

In our present work we introduced a new position based clustering approach as an enhancement of the existing XGAF algorithm which combines the advantages of clustering and covering and gets the best of both, i.e. full coverage and high lifetime. Hence our algorithm introduces a new abstraction level by bringing the idea of temporary deactivation up to the cluster level.

The intercluster rotation of our approach is described as a two phase day-night-cycle. The intracluster rotation, however, is left open and may be realized by any common algorithm.

Additionally, we introduce a self-healing algorithm as an aspect of clustering algorithms. As shown in section V, the modified cell size in combination with our day-night-cycle as well as the assistance algorithm leads to a longer lifetime, higher node efficiency and a higher coverage over time, which also leads to a high amount of robustness.

Although we investigated a large number of aspects, several points have been left open, which will be investigated in future work. An essential question will be the networks behavior in case of imprecise localization, which represents the real conditions up to now. It is also necessary to have a closer look on different cluster sizes to find out which is the most effective size in respect of energy-awareness, number of required nodes and robustness.

References

[1] Salzmann, J.; Kubisch, S.; Reichenbach, F.; Timmermann, D.; "Energy and Coverage Aware Routing Algorithm in Self Organized Sensor Networks", Fourth International Conference on Networked Sensing Systems; Braunschweig, Germany, 2007

[2] Abrams, Z.; Goel, A.; Plotkin, S.; "Set k-cover Algorithms for Energy Efficient Monitoring in Wireless Sensor Networks, "Proceedings of Information Processing in Sensor Networks 2004, Berkeley, USA, 2004

[3] Liu, C.; Wu, K.; King, V.; "Randomized Coverage-Preserving Scheduling for Wireless Sensor Networks, "4th International Networking Conference, Waterloo, Canada, 2005

[4] Heinzelman, W. R.; Chandrakasan, A.; Balakrishnan, H.; „Energy-Efficient Communication Protocol for Wireless Microsensor Networks", IEEE Proc. Hawaii Int'l. Conf. Sys. Sci., Jan. 2000

[5] Xu, Y.; Heidemann, J.; Estrin, D.; "Geography-informed Energy conservation for Ad Hoc Routing", Proceedings of the 7th annual international conference on Mobile computing and networking, Rome, 2001

[6] Inagaki, T.; Ishihara, S.; "A Scheme for Expanding Grid Size of Geographical Adaptive Fidelity", Fourth International Conference on Networked Sensing Systems; Braunschweig, Germany, 2007

[7] Schurgers, C.; Srivastava, M. B.; "Energy Efficient Routing in Wireless Sensor Networks", Military Communications Conference, 2001

[8] Hong, X.; Gerla, M.; Wang, H.; "Load Balanced, Energy-Aware Communications for Mars Sensor Networks", Proceedings of Aerospace Conference, 2002

[9] Ma, M.; Yang, Y.; "Adaptive Triangluar Deployment Algorithm for Unattended Mobile Sensor Networks", IEEE Transactions on Computers, Vol.56, No.7, July 2007