An Ecologically Inspired Intelligent Agent Assisted Wireless Sensor Network for Data Reconstruction

Fan Bai, Kumudu S. Munasinghe, and Abbas Jamalipour
School of Electrical and Information Engineering
University of Sydney, NSW 2006, Australia
{fan, kumudu, abbas}@ee.usyd.edu.au

***Abstract*—One of the most important problems studied in data harvesting wireless sensor networks (WSNs) is the optimization of the tradeoff between the accuracy of the reconstructed field data and the resource consumption. In order to optimize the resource consumption, whilst not compromising the accuracy of the reconstructed field data, an ecologically inspired marginal value theorem strategy (MVTS) is proposed for a mobile agent for choosing the next sensor node to be visited in the data acquisition process. The proposed MVTS can adaptively gain new knowledge during the process of collecting observations from a WSN comprising of static sensor nodes. Therefore, only the relatively important sensor observations will be collected by the agent according to the variety of the background environmental data. This is thought as an efficient way to reserve the resources, such as energy and bandwidth, because only the important observations are collected. Illustrated analytical and simulation results confirm the above achievements.**

I. Introduction

In the recent years, wireless sensor networks (WSNs) have been applied in a wide range of domains for data harvesting, monitoring and so on. Regardless of its application domain, a majority of end users are interested in reconstructing an accurate picture of the field with the use of collected observations from sensors. In traditional WSNs, observations are relayed by sensors to sink nodes. However, this inevitably leads to higher overheads due to complicated data routing mechanisms, large amounts of data exchange, and hot spot problems on relatively weaker sensor nodes. Therefore, to overcome these aforementioned problems, a viable solution is presented by the data MULE approach [1]. The MULE is a mobile agent which moves within the sensor field for directly collecting observations from sensors. Different flavors of mobile agents assisted WSNs are investigated by researchers and engineers, such as [2], [3].

However, almost all of the currently available approaches collect observations from all sensor nodes of the WSN, which overlooks the correlation and redundancy that exists among them. Real life environmental data may show certain levels of spatial correlations in particular areas. In the high correlated areas (quiet areas), the data changes gradually along the spatial domain, while in the low correlated areas (busy areas), sudden changes may be observed. In the area of data aggregation techniques for WSNs, exploiting the correlation is a popular method for removing the redundancy. However, this category of approach requires each sensor node to transmit at least one packet to the aggregation center.

This work is partly supported by the Australian Research Council under the Discovery Project Scheme (DP1096276).

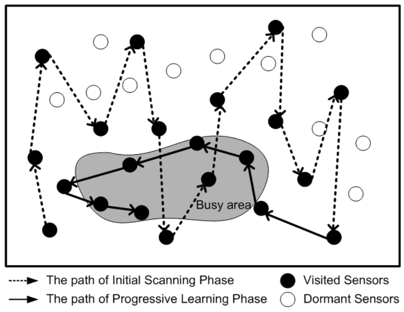

Figure 1. The working process of proposed sensor visiting strategy.

In this work, an alternative approach to solve this problem is presented. We use an intelligent mobile agent, which is embedded with marginal value theorem strategy (MVTS) to roam inside the sensor field for collecting observations from static sensors. The idea of MVT is inspired through ecological behavior. The basic principle of MVT conceptualizes how an optimally foraging animal exploits the resources distributed in various food patches [4]. Hence, the MVT can be applied for gaining that relatively sufficient benefit by spending the minimal cost in a way the rate of gain is maximized. The mobile agent is able to adaptively detect quiet/busy areas and adjust the corresponding sampling rate by choosing to visit the relatively important sensors. The entire process consists of two phases. In the initial scanning phase, a rough knowledge is learnt by the mobile agent, following which the progressive learning phase drives the mobile agent to visit those relatively important sensors. An overview of the proposed framework is illustrated in Fig. 1. In comparison to other mobile agent assisted WSNs, the potential advantages of our proposal are:

- The total resource (energy, bandwidth) consumption is reduced by ignoring some sensors in quiet areas. However, the visiting rate of sensor nodes in the busy areas is relatively higher to guarantee the accuracy.
- The latency for the mobile agent to complete a round of collection may be reduced via bypassing dormant sensors.

The remainder of this paper is organized as follows. Section II provides an overview on the related works and the state-of-the art. Section III, outlines the system structure and assumptions followed by the description of the marginal value theorem strategy under Section IV. Section V presents some examples and simulation results to demonstrate the correctness and efficiency of MVTS followed by the concluding remarks.

II. RELATED WORKS

In current literature, concept of mobility is applied to WSNs in various forms. Its first successful application on a real life project is ZebraNet, where data is collected from zebras with attached sensors by exploiting the natural motion of animals [5]. Alternatively, attempts are made for intermittently activating randomly picked up sink nodes from a WSN [6]. According to [6], this method evenly drains the energy of the sensor nodes, thus increasing the lifetime of the WSN. Also a different flavor of the above method is used in [7], where mobility is introduced to the sink node. However, due to the exchange of sink re-location updates, the overheads incurred in these approaches remain high.

Therefore, in order to avoid such overheads associated with multi-hop routing, a viable solution is presented by the data MULE approach [1]. The data MULE approach uses randomly roaming data harvesting agents for directly collecting data from individual sensors. However, the disadvantages of this architecture are high latency and best-effort-delivery. In light of improving this, interest is shifted towards using data collecting agents with controlled mobility. For example, [8] propose controlled mobility models for data access with built-in data aggregation methods in place, in which only a selective number of nodes are accessed by the data collecting agent, hence considerably reducing latency. There are also other approaches where stop-stations similar to cluster heads are visited for data collection in a WSN [9]. However, those methods require a certain level information exchange among sensor nodes, which eventually consumes high energy.

On the other hand, mobile agents can be programmed with predefined trajectories, especially for applications related to data harvesting [10]. Similarly, there are works on predefined scheduling algorithms proposed for such mobile data collecting agents [2]. Therefore it can be summarized that the common goal of the above works is to optimize the latency-energy trade-off as systematically investigated in [3]. Nevertheless, all the aforementioned works require data harvesting agents to visit all sensors in order to maximize the amount of obtained information. Hence the redundancy of the data is also not removed by these collecting strategies.

The motivation for developing a sophisticated collecting strategy for intelligent mobile data collecting agent is to balance the tradeoff between accuracy and resource consumption. The proposed MVTS is able to exploit the spatial correlation, and remove the redundancy by only visiting those sensors with relatively important observations. Similar thinking is found in [11], where the authors developed a special media access control (CC-MAC) protocol for preventing every sensor node reporting their observations. However, the CC-MAC is used in traditional WSNs for applications like event detection. The following proposed MVTS does not put any heavy load on weal sensors, and can be perfectly applied to field data reconstructing applications.

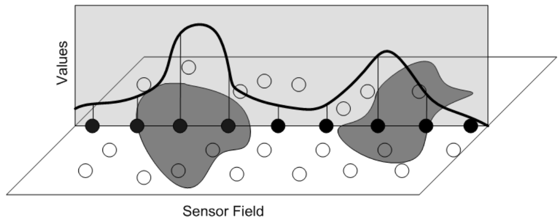

Figure 2. A slice view of two dimensional sensor field and the corrsponding values.

III. SYSTEM MODEL

This section presents the system model underlying the MVTS method, and describes the assumptions made within reasonable limits. The entire system consists of a large number of normal sensor nodes and an intelligent mobile agent.

Each sensor node has limited power and processing capability. It is inconvenient or unfeasible to replace or recharge the sensor nodes in most practical sensors. The sensor nodes are assumed to be static and identical. For the purpose of field data reconstruction, in each sampling round, the sensor nodes are responsible for reading the environment synchronously and storing the values. The radio model of sensor node works in a reactive way, which means that one sensor node passes its current observation to the intelligent mobile agent only when the mobile agent is visiting it. For the energy saving purpose, the transmission range is assumed to be relatively small. It is almost good to say that the sensor node can communicate with mobile agent when they are in the same location.

The mobile agent is equipped with an ecologically inspired intelligence, which can roam inside the sensor field to collect observations from sensor nodes. It is assumed that the mobile agent is a robot or a vehicle with controlled motion. Therefore, it is also reasonable to assume that the mobile agent is able to record all of the locations of sensor nodes. The mobile agent is relatively more powerful than a sensor node. It is easily to be recharged. Once the mobile agent visits a specific sensor node, it collects and stores the observation from the sensor node. Furthermore, upon the fresh information learnt from a new observation, the mobile agent uses MVTS to choose the next sensor node to be visited. Because of the transmission time of a message is negligible with respect to the time it takes the agent to meet a sensor node, the transmission may only take place at meeting times and are assumed to be instantaneous.

In order to maintain clarity and convenience for explaining the MVTS, a simplified one dimensional sensor field is used in this paper. This scenario could represent a slice view of a normal two dimensional field as illustrated in Fig. 2.

IV. MARGINAL VALUE THEOREM STRATEGY (MVTS)

Marginal value theorem is considered as the optimal foraging behavior of an animal for exploiting the food resources distributed in patches [4]. The foraging animal uses the rate of gain of marginal value to decide when to leave the patch and start searching for a fresh one. In the case of mobile agent assisted WSNs, each observation gathered by sensor node contains certain amount of information, which can be used to reconstruct the field data. Because of the spatial correlation and redundancy, the relative information gain after removing the redundancy from an observation can be mapped as the marginal value, and the corresponding sensor node can be mapped as the patch. The intelligent agent can be equipped with MVTS to exploit the sensor nodes and complete the job.

In order to apply MVTS in the WSN, the following requirements must be met:

1) The amount of information contained in a single independent observation is identical.

2) The relative information gain of a new collecting observation is the amount of information increased after the redundancy is removed.

3) The information amount contained in a group of observations should be the information amount after the redundancy is removed.

A. Observation Vector Quantization and Marginal Gain

An observation consists of the location of the corresponding sensor and the value of the gathered data. Hence, the observation is represented as a vector <location, value>. Therefore, the observation space is made by spatial domain and value domain. As shown in Fig. 3, an observation is a dot in the sensor field and value coordinates.

Observation vector quantization works by dividing a large set of observation vectors into groups, which have approximately similar properties based on the user's resolution requirements. Each group is represented by its centroid vector. So, as we can see from Fig. 3(1), an observation A represents other observations around it, which is grouped by a rectangle. The information contained in an observation is represented by the area of the rectangle. Hence, the requirement 1 is satisfied.

In Fig. 3(2), a new observation B is collected, in which an identical amount of information surrounded by a rectangle. However, the new rectangle B is partially overlapped (gray area) with the previous rectangle which belongs to A. The overlapped part is considered as the redundancy between observation A and B. The relative information gain from observation B is the area after removing the redundancy. It is shown as the blank area of the rectangle. Same principle is applied to observations C and D in Fig. 3(3) and Fig. 3(4). Hence, the requirement 2 is satisfied.

The total information amount of a group of observations is the total area covered by these four rectangles in Fig. 3(4). The overlapped area only counted once, where the redundancy is removed. Hence, the requirement 3 is satisfied.

B. Decision Module

The MVTS is applied in the decision module, which decides the actions of the mobile agent. As mentioned in the Section IV.A, the precision of the observation vector quantization is based on the resolution requirements. The resolution requirements contain the following metric:

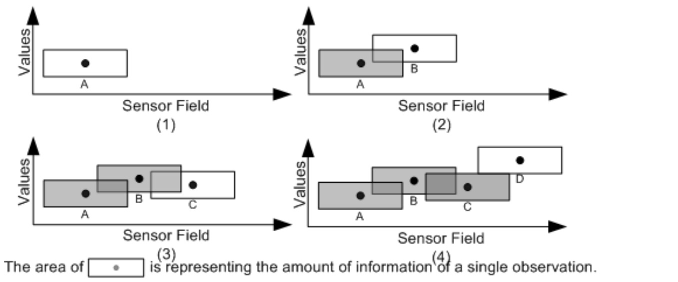

Figure 3. The vector quantization system of observations and the relative information gain process of an observation.

● *Spatial Precision*: It is specified by the user to decide the spatial range of a group of observation vectors, which can be represented by the centroid vector of that group. This is visually represented by the width of an observation rectangle in Fig. 3. The unit of *Spatial Precision* is meter.

● *Value Precision*: It is specified by the user to decide the value range of a group of observation vectors, which can be represented by the centroid vector of that group. This is visually represented by the height of an observation rectangle in Fig. 3. The unit of *Value Precision* is same as the unit of the sensed environmental data.

The decision module works in two phases, the initial scanning phase and the progressive learning phase. In the initial scanning phase, the mobile agent visits a subset of sensor nodes, which are separated by the spatial precision. This ensures that each observation collected from sensor nodes contributes 100% of its information for reconstructing field data. At the end of the initial scanning phase, it also gives a rough knowledge of the environmental data of the sensor field.

In the progressive leaning phase, the mobile agent learns the new knowledge from each new observation and decides the next sensor will be visited, which is expected to bring the highest rate of the relative information gain. Other rules that MVTS must comply with are listed as:

1) The expected relative information gain from the new observation must be larger than a predefined threshold, *TH*. For example, 50% of the information amount contained in an independent observation.

2) The cost of taking the next observation needs to be minimized. The cost of taking an observation consists of two parts. The first part is the energy consumption, which takes place when the messages are exchanged between the sensor node and the mobile agent. This part of cost is identical for every sensor nodes. The second part is the resource consumption of the mobile agent while traveling from one place to another place. This part of cost is proportional to the distance between two places. It is different from case to case.

To strive on these rules, at the beginning of the second phase and after collecting each new observation, the mobile agent needs to complete several steps.

Step 1: The mobile agent will estimate the observations gathered by the unvisited sensor nodes. The estimated observation values are obtained by interpolating techniques, such as Spline interpolation. The differences of estimation between various interpolation techniques are negligible, as long as the reasonable trend of data change is predicted.

Step 2: The mobile agent needs to find out a Candidate Observation Set, which all the estimated observations inside this set are expected to relatively contribute more than TH percentage of its original independent observation on top of the already obtained observations. Then the mobile agent sorts these estimated observations according their corresponding sensor locations. If this set is empty, the mobile agent will stop visiting new sensors and finish the entire process.

For example, in Fig. 4(1), suppose the TH is 50%, the observations A and H are those which have already been collected by the mobile agent, and the current location of the mobile agent is at the location of the observation A. The sorted Candidate Observation Set is {B, C, D, E, F, G}. Every observation in this set is expected to provide more than TH percentage of its information on top of the information contained in observations A and H.

Step 3: The mobile agent checks each observation inside Candidate Observation Set one by one to produce the Relative Information Gain Set. In each round, the percentage of new relative information gain corresponding to the currently checked observation is added to the Relative Information Gain Set, and the previous elements are adjusted. The adjustment is made by checking if the rectangle corresponding to the being checked observation is redundant with other observations. The complexity of this algorithm lies in track the relationship of each observation to make the correct adjustment. The result of this step is the Relative Information Gain Set, which contains the relative information gain of each candidate observations after remove the redundancy.

In Fig. 4(1), the Relative Information Gain Set for the Candidate Observations is {45%, 40%, 90%, 30%, 50%, 50%}.

Step 4: The mobile agent checks the Relative Information Gain Set, the first element of this set that is not smaller than TH is chosen, and the sensor of the corresponding observation in the Candidate Observation Set is the next sensor going be visited. After new observation is collected, the process returns back to step 1. The new observation is deemed as the qualified observation with minimized cost of collection.

In Fig. 4(1), the element which satisfies the condition is the third element (90%), and the corresponding observation is D. The mobile agent is going to visit the sensor which the estimated observation of this sensor is D next. The mobile agent will finally collect observations D and F, in Fig. 4(2). As it can be noticed, all the final collected observations relatively contribute more than 50% of its contained information to the application. On the other hand, any other observations from other sensors will not contribute more than 50% of its contained information to the application.

The proposed strategy described above is able to drive the mobile agent to visit more sensor nodes in busy areas. Hence, the sampling rate in those areas is higher than those in quiet areas. This efficiency of MVTS is verified in the next section.

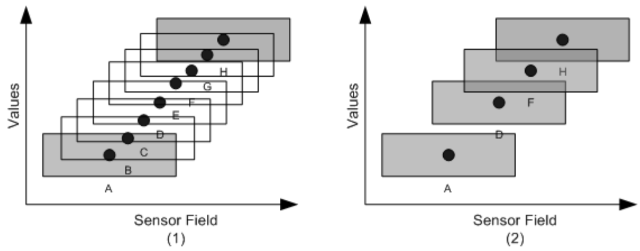

Figure 4. An example of the mobile agent use MVTS to choose the sensors to be visitied.

V. PERFORMANCE EVALUATION

In this section, our proposed MVTS is evaluated on a one dimension sensor field on Matlab. In the simulation, 100 sensor nodes are evenly distributed on a 1000 meter long field. The background environmental data are assumed to have a certain level of correlation. The correlation function used to produce the background data follows an exponential model [11]:

$$corr_{i,j} = \exp(-d_{i,j} / \theta) \quad (1)$$

The i, j are two locations in the sensor field, and $d_{i,j}$ is the distance between these two locations. The constant θ is the correlation factor and is set to 30 in the simulations, which represents a general level of correlation. Using equation (1), 1000 correlated Gaussian random variables are produced as the background environmental data. The simulation reconstructs 1000 values for every scenario. The Value Precision is preset to 1, and the threshold, TH, is preset to 50%. The mobile agent is assumed travels in a constant speed of 1 meter per unit time.

In Fig. 5(1), the relationship between the Spatial Precision and the number of visited sensor nodes by mobile agent is displayed. As it can be seen from the figure, as the value of Spatial Precision increased, the number of visited sensor nodes decreased. This is because an increase in the value of Spatial Precision actually decreases the precision and the requirement of sampling rate is also decreased. It requires that all 100 sensor nodes are visited when the Spatial Precision is set to 10 meters, and requires only 54 sensor nodes to be visited when the Spatial Precision is set to 100 meters. However, it is worth to note that in the first half of this plot, the number of visited sensor nodes drops dramatically, while, when it moves to the second half of the plot, the dropping speed decreases rapidly.

In Fig. 5(2), the relationship between the Spatial Precision and the reconstructed distortion is displayed. The reconstructed distortion is defined as follows,

$$\frac{\sum_{i=1}^{1000} \left| \textit{The Reconstructed Value}_i - \textit{The Real Background Value}_i \right|}{1000} \quad (2)$$

Corresponding to Fig. 5(1), when the Spatial Precision is set to 10 meters, all 100 sensor nodes are visited, the reconstructed distortion is minimal, which is equal to 0.543. However, as the value of the Spatial Precision increased (actual precision decreased), the distortion is increased. It also needs to mention about that the increasing speed of the second half of distortion plot is faster than the first half.

Jointly considering the plots in Fig. 5(1) and Fig. 5(2), from an engineering prospect, the trade-off between the accuracy and resource consumption needs to be optimized. The value of Spatial Precision is chosen to be between 40 meters to 50 meters, which will bring the maximal benefit to the system.

In Fig. 6, the detailed information is demonstrated for Spatial Precision is set to 50 meters. It includes the reconstructed data, the visited sensor nodes and the visiting schedule. As seen from Fig. 6(1), the reconstructed data are attached to the trend of real environmental data, and the visited rates of sensor nodes in different areas vary. The visited rates in busy areas are relatively higher than those in quiet areas. In this scenario, a total of 63 sensor nodes out of 100 have been visited by mobile agent. Fig. 6(2) displays the visiting schedule of mobile agent. From 0 to 1000 time units, the mobile agent works in the initial scanning phase. It visits sensors, which are separated 50 meters away from each other. From 1000 time units to the end of the process, the mobile agent works in the progressive learning phase. It decides the next sensor node going to be visited based on the rate of information gain. It can be noted that, there are some backward steps in second phase. That is because the new information makes the mobile agent adjust its previous estimation. However, the scheduling is still very efficient. It only takes 2403 time units to finish the entire process.

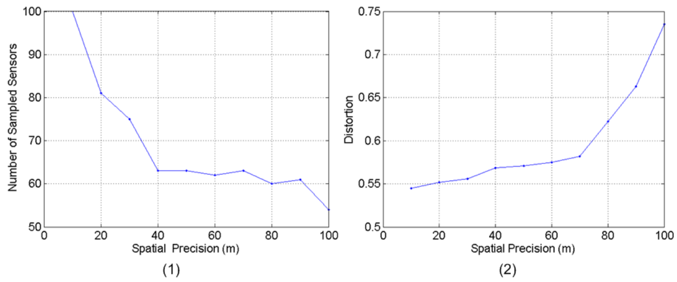

Figure 5. The number of visited sensor nodes vs. the *Spatial Precision*.

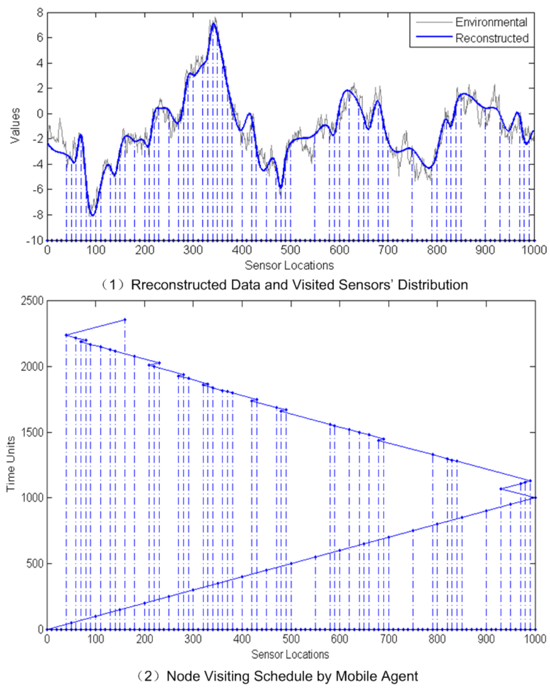

Figure 6. Detailed information when *Spatial Precision* is equal to 50 meters.

VI. CONCLUSIONS

In this paper, the problem of cross optimization for accuracy and resource consumption in field data reconstruction in WSNs using intelligent mobile agent was investigated. The novelty of our solution was based on the use of the previously unexplored marginal value theorem strategy, which decides the next sensor node going to be visited based on progressively learning the new information from the sensor field. Therefore, the accuracy and resource consumption in a WSN was optimized by adaptively increasing the visiting rate of sensor nodes in relatively busy areas. In order to illustrate the advantages of the proposed procedure, the reconstructed data and visited sensor nodes distribution are displayed; the feasibility and the validity of the proposed solution were demonstrated. The future works include extend the sensor field to a general two dimension field (length and depth), which is deemed to bring more advantages, such as reducing the latency of colleting observations via bypass those sensors which can be ignored.

REFERENCES

[1] S. JAIN et al., “Exploting Mobility for Energy Efficient Data Collection in Wireless Sensor Networks,” *ACM Mobile Networks and Applications*, vol. 11, no. 3, pp. 327-339, 2006.

[2] Y.Gu et al, “Data harvesting with mobile elements in wireless sensor networks,” *ScienceDirect Computer Networks*, vol. 50, no. 17, pp. 3449-3465, 2006.

[3] R. Sugihara and R. K. Gupta, “Optimising Energy-Latancy Trade-off in Sensor Networks with Controlled Mobility,” in *Proc. IEEE Conference on Computer Communications (INFOCOM)* , pp. 2566-2570, 2009.

[4] E. L. Charnov, “Optimal foraging: the marginal value theorem,” *Theoretical population biology*, vol. 9, pp. 129-136, 1976.

[5] S. JAIN et al., “Energy-Efficient Computing for Wildlife Tracking: Desgin Tradeoffs and Early Experiences with ZebraNet,” *ACM SIGARCH Computer Architecture News*, vol. 30, no. 5, pp. 96-107, 2002.

[6] S. R. Gandham et al., “Energy Efficien Schemes for Wireless Sensor Networks with Multiple Mobile Base Stations,” in *Proc. IEEE Global Communication Conference (GLOBECOM)*, vol. 1, pp. 377-381, 2003.

[7] J. Luo and J-P. Hubaux, “Joint Mobility and Routing for Lifetime Elongation in Wireless Sensor Networks,” In Proc. *IEEE Conference on Computer Communications (INFOCOM)*, vol. 3, pp. 1735-1746, 2005.

[8] N. Nguyen et al., “Improving Information Availability in Storage-Centric Sensor Networks,” in *Proc. IEEE Conference on LocalComputer Networks (LCN)*, pp. 82-90, 2008.

[9] H. M. Almasaeid and A. E. Kamal, “Modeling Mobility-Assisted Data Collection in Wireless Sensor Networks,” in *Proc. IEEE Global Communication Conference (GLOBECOM)*, 2008.

[10] A. Srinivasan and J. Wu, “TRACK: A Novel Connected Domination Set based Sink Mobility Model for WSNs,” in *Proc. IEEE International Conference on Computer Communications and Networks (ICCCN)*, pp. 1-8, 2008.

[11] M. C. Vuran and I. F. Akyildiz, “Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks,” *IEEE Tran. on Networking*, vol. 14, no. 2, pp. 316-329, 2006.