Tree-Based Organization for Very Large Scale Sensor Networks

Moran Feldman Faculty of C.S. Technion Israel Institute of Technology Haifa, Israel moranfe@technion.ac.il

Abstract— One of the major challenges in deploying large scale sensor networks is the ability of the sensors to weave dynamically and autonomously into a sensing plan. In this paper we present a novel algorithm with certain characteristics. It is applicable for thousands of sensors and uses a tree based organization to present an aggregation method that aggregates discrete events into compound ones. It presents a set of security levels to ensure that events are transmitted to the control center (sink) and also presents backup layers to ensure maximal connectivity. Furthermore, it is applicable for sensors with no GPS. The algorithm was successfully tested using the dedicated simulator on a terrain containing 10,000 sensors. Our results show that the sensors perform the process of weaving into a sensing plan, the task of identifying multiple intruders, reporting the events to the sink in a short time and comply with the other demands.

Keywords- Large Scales Sensor network; Routing; Localization; Data aggregation.

I. INTRODUCTION

The evolution of microelectronics and communication technologies facilitates the manufacturing of miniature sensors comprising a small transmitter/receiver, a processor, memory components and a low-power battery [1], [2]. Most often, the sensor, or node, is a Boolean sensing device that detects an event within a given sensing range, and is able to inter-communicate using wireless protocols with adjacent nodes, creating a wireless sensor network (WSN).

There are a few common methods to categorize the network organization in a WSN according to the network structure [3]. The first category is flat routing, wherein all nodes have an identical role. The routing of events from the sensing nodes to the sinks can use any node in the network without any limitations and without interacting with any centric nodes [4][5][6][7][8][9]. In hierarchical routing protocols, part of the sensors possesses additional tasks. In this case, the sensors are grouped into clusters. One of the sensors in every cluster is designated as the cluster-head. If needed, it is possible to group cluster heads into new clusters [10][11] [12][13]. This organization method delegates routing responsibilities to the cluster heads rather than the regular nodes. Hierarchical routing is considered as an efficient way to reduce energy consumption by transferring the aggregation process to the cluster head. A third method is

Sharoni Feldman

ATC Consulting Haifa, Israel feldsh@netvision.net.il

location-based routing. The sensors are aware of their position in the theater. This location information can be exploited in order to route data over the network more efficiently. The sensor location can be obtained from a GPS receiver installed in every sensor [14]. Another method uses relative coordinates that are based on information gathered from neighboring sensors. The distance between neighboring sensors can be estimated according to the strength of the incoming signals [15][16][17]. An interesting approach is presented in [18]. The traffic is divided into two types: high priority traffic and low-priority traffic. High priority data is routed using a dedicated congestion zone arranged as a spanning tree with the sink as a root and the low priority traffic is routed via other nodes and longer paths. This model enables every node to be connected to several trees according to the number of sinks.

In this paper, we present the Very Large Scale Sensor Network Algorithm (VLSSNA). The VLSSNA presents a novel routing technique for a very large network composed of 10,000 sensors. This number is significantly higher than the number presented in literature [19]. This algorithm creates a flat network. Moreover, the algorithm enables the network operators to assign an "importance level" to events and to transmit important events on parallel layers of the network. This feature increases significantly the probability that events will not get lost even if some intermediate nodes malfunction.

The sensors dissemination process is totally random without any prior definitions in the sensors. The organization of the network is an autonomous procedure without any external intervention. In our research, we assume a sensor density that prevents an object from crossing the sensing field without being detected [20]. Energy economization is achieved using a communication method that combines broadcasting and sensor-to-sensor communication and an aggregation process used to minimize the number of messages.

The reminder of this paper is organized as follows. Section 2 presents the theater and the network elements, Section 3 presents the sensors control and management, Section 4 deals with energy saving methods, Section 5 presents the simulator used for evaluation of the sensor network and Section 6 presents the simulations and results. Conclusions follow in Section 7.

II. THEATRE AND NETWORK ELEMENTS

A. The Theater

Fig. 1 presents a schematic view of a typical theater on which the sensors are dispersed. The sensors are dispersed so that their density ensures that every target will be detected by one or more sensors. At the edge of the theater, three base stations (BSs) are placed, arranged in an equilateral triangle. The BSs are connected by a high-speed communication link. At least one sink controls and monitors the events.

B. Network Elements

a)Base Stations

The BSs are identical, and they are controlled by the sink. Every BS is required to know its exact position in space (x, y, z) coordinates. This information can be obtained manually when the BS is installed or via a GPS. In addition, all BSs clocks are synchronized.

A BS is constructed of the following units: (a) A long-range downstream transmitter that covers the whole theatre. (b) A short-range downstream transmitter and upstream receiver that cover the adjacent sensors. (c) A high speed LAN that connects all BSs and the sink. This LAN is used to transfer synchronization data among the BSs and the sink, alarms received from the sensors to the sinks and messages from the sink to the BSs and the sensors.

b) Sensors

The sensors are scattered in the terrain. The distribution of the sensors in the theater is not required to be uniform. However, it is assumed that some connectivity between a sensor and its surrounding sensors can always be found. A sensor has a limited life expectancy, and as a dispensable device the whole network will continue to function with a certain decrease in the number of active sensors. The design of the network allows the operators to add periodically new sensors to the field to overcome the natural decrease in the number of sensors. A sensor is composed of the following major units: (a) A binary Omni sensing device like a microphone. The sensor is unable to detect the direction of the event. (b) Short range Transmitter/Receiver. We can assume that the transmission range is greater than the sensing range. Fig. 1 presents a sensors field. The dashed lines present the sensing range while the full lines present the transmission range. A sensor can receive messages transmitted from other sensors within their transmission range or from the BSs. (c) Processor and Battery.

III. SENSORS CONTROL AND MANAGEMENT

A. First Network Activation

The network activation process starts after completing scattering the sensors in the field. The sensors in the field will not start to work until the activation process ends. Note that all sensors clocks will be synchronized during the activation process. The activation process has 3 steps and is coordinated by the BSs:

1. BS₁ sends a beacon that covers the whole field. This beacon carries the following data elements: (a) Time stamp used to synchronize the internal clock of every sensor and the other BSs. (b) The name and geographical location (XBS₁, YBS₁, ZBS₁) of BS₁ (c) Wait time (t_w) in milliseconds.

Every sensor receiving this beacon will update its internal clock with the BS_1 time and its internal database with the coordinates of BS_1 . The other BSs will update their internal clock according to BS_1 clock.



Figure 1: A typical sensor theater. Three base stations organized as an equilateral triangle and a single sink monitor the wireless sensor network.

- 2. BS₂ waits predefined t_w milliseconds (a) value known to all network elements) before broadcasting a beacon. This beacon which covers the whole field carries the following parameters: (a) The name and geographical location (XBS₂, YBS₂, ZBS₂) of BS₂. (b)Time stamp. Every sensor receiving this beacon will update its internal database with the coordinates of BS₂.
- 3. BS₃ will wait t_w milliseconds after BS₂ transmission before broadcasting a beacon. This beacon parameters and the process are identical to these of BS₂. Every sensor receiving this beacon will update its internal storage with the coordinates of BS₃.

After receiving the 3 beacons, every sensor starts its localization algorithm (based on trilateration). The termination phase of this algorithm is a set of sensors spread in the field when every sensor has identified the BSs, the BSs geographical locations and its own location. In the current state, the sensors stay dormant and do not initiate any activity.

The network activation process runs periodically and is used to join new sensors that were added to the field or remap old sensors that were moved inside the field.

B. Messages and Data transfer

The communication among the network elements is performed by messages. The network elements exchange broadcast messages addressed by all listening nodes within the transmission range and directed messages that address a specific node within the transmission range. Another type of classification is based on the transmission range of the message originator. A BS can transmit short range and long range messages while a sensor is capable to transmit only short range messages. Tab. 1 summarizes combinations between the message types and message transmission range. While a BS can send directly a long distance message to every node in the terrain, a node which is required to send an event to the BS, is required to use intermediate nodes to bridge the distance. The process of transferring the information from the node to the BS is based on a "store and forward" mechanism. A node that received a message will forward the message to the next leg in the chain only after it was received completely.

Special attention was given to energy saving. As will be described in details later in this paper we implemented "energy saving" methods in critical and demanding procedures.

C. Trees formation processes

During the trees formation phase, every sensor builds its connections in the field. All BSs initiate stimulatingly the trees formation process although the processes are independent. It starts after the localization process and terminates when the sensors complete joining the spanning trees. After completion, every node is connected to n trees when n is the number of BSs.

Fig. 2 presents 2 trees in the theater. The tree of the root node BS_1 and the nodes addresses is presented in continuous lines in black and the tree of root node BS_3 is presented in dashed lines. The addresses of the nodes that belong to root BS_1 start with <1.> and the addresses of nodes that belong to BS_3 start with <3.>. We skip the tree formation process due to lack of space.

IV. DETAILED DESCRIPTION - TFA ALGORITHM

The trees formation algorithm (TFA) organizes the nodes in the field in rooted trees. Only nodes that belong to the same tree can transfer events to the BS which is also the tree root and the sink. To ensure the maximal connectivity, all nodes will try to organize themselves in a single tree. Every node in the field has a unique and fixed node-id and a virtual coordinate (of the type x.y.z..) that may change depending on the changes in the tree structure. Every tree is identified by a

MSG Range Type	Long Range Message	Short Range Message
Directed Message	directed to a specific sensor	1
Broadcast Message	A message transmitted only by a BS. This message is targeted to all sensors in the terrain	A message sent by sensors or the short-range transmitter of the BS. All receiving nodes within the short transmission range address this message.

Table 1: Messages Definition and Types of Messages

"tree name" which is the *id* of the root node, which is the BS. The metrical join (referred also as join) protocol should satisfy the following properties:

Eventually all nodes within transmission area must fuse into a single tree. When two trees are being fused, most updates should be made to the nodes of the smaller tree (in the number of nodes). The protocol should maximize the number of nodes that joins the tree in every step (yielding a parallel fuse). Nodes periodically attempt to balance the tree by shortening their distance to the root of the tree by improving their position in the tree and joining higher-level nodes.

The protocol is fully distributive with no "central" control.

Running alarms will not be affected (i.e. will not break) in any way from the ongoing trees join process or rearrangement of an existing tress. The parallel process of creating the trees (with the roots BS_1 , BS_2 ...) is fully independent. When all processes terminate, every node is a member in all trees.

The TFA algorithm runs when the stability of the network is violated. Every time this algorithm runs, it will bridge the "holes" created by faulty sensors and will add new sensors that were added to the terrain.

A. Events, Events aggregation and Delivery Verification

In our model, every sensor has a circular events detection area around itself. An intruder that enters the sensors field will stimulate every node whenever it enters its detection radius.

Every node that detects an intruder sends an event message to its father. Fig. 3 presents a section of the sensors field and the intruder path within it. The circle around each sensor presents the detection zone of the sensor. Our assumption is that every sensor is able to perform basic filtering of the detected noise, i.e. a sensor programmed to detect a noise of a car will not respond to a noise created by a walking animal



Figure 2: Trees in the terrain

or barking of a dog. In Fig. 3 the intruder triggers 4 sensors located on his path.

The tree structure creates a natural organization of the nodes, which allows the network to "aggregate" events from each subtree in the subtree root and minimize the amount of events transferred towards the sink. For example, node <1.1> is able to aggregate the events detected by its children <1.1.X> and send it as a unified event to its father. A

managing automaton runs in every node and enables the node to act in one of two possible ways: Detect an event or receive and handle event messages from its children and transfer them either transparently or with some updates toward the sink.

The automaton is composed of 3 states.

1) Idle. This is the stable state of the node. In this state, the node is waiting for an event created by the intruder or to a message from one of its children. When the node detects an intruder, it sends the message "event-reportmessage" with the event details to its father. The main event details are the geographical location of the node and the event timestamp. After detecting an event, the node will not report any additional event for a short time span. This prevents the node from creating a flood of messages toward the sink caused by a single short timed stimulation.

2) Store information – open timer. As soon as the node receives an "event-report-message" from one of its children, it assumes that his other children may detect this same event. The node stores the event data and waits for additional messages from the other children. A short wait timer is activated in order to limit the wait time. If the node receives an "event-report-message" from all its children, the timer is redundant, and it is cancelled.

3) Calculate direction. This state is activated once the node decides to stop waiting for additional "event-report-messages". The node tries to calculate the local direction of the intruder, based on the messages it received. The results are then sent towards the root using the message "aggregated-event-message". It is possible to calculate the direction only if two or more children reported the event. In case that only a single child detected the event, the aggregated message will carry the location of the detecting node. In the case that 2 or more nodes have detected the event, the message will carry the locations of the two most distant detecting nodes.



When a node receives a message "aggregate-eventmessage" from its subtree, it acts as a router and sends it to its father.

B. Data Aggregation

The purpose of the aggregation process is to reduce the amount of data transferred from the detecting nodes toward the root. However, as will be explained later, this process slows the speed in which events are transferred toward the sink. The aggregation process can be nested, for example, node <1> in Fig. 3 can aggregate the data from its children <1.1> and <1.2> into a single message. A *k-level* aggregation process performs the aggregation process in the *k* lowest levels of the tree (starting from the leaves). The aggregation level can be adjusted according to the required performance of the network and the type of expected intruders

The need to aggregate messages requires that inbound messages will be delayed in the node for a period of time. This delay enables other inbound event messages to arrive during the delay and to be aggregated into a single aggregated message.

C. Event Delivery Verification

The verification feature enables the network to ensure with high level of certainty that a reported event has been transferred successfully from the sensing nodes via the routing nodes to the BSs and the sink. The network architecture enables the use of two types of verification levels. The basic method is "Report & Forget" which does not require a BS to acknowledge the acceptance of the event report. The enhanced method "Report & Acknowledge" requires the receiving BS to send an acknowledgement to the reporting nodes. The acknowledgement does not use the sensors network to transfer the acknowledgement to the sender but broadcasts a "Shout" message over the air from the BS with the sensor-id and event details. In case that the event originator has not received the acknowledgment message within a predefined period, it will resend the event report.

Another capability of the architecture is to use two levels of energy saving delivery methods. The economized method uses a "*Load-Sharing*" mechanism where the reporting node selects cyclically one of the trees and sends the event report over the tree to a single BS. The "*Active-All*" mechanism is more energy consuming. Using this method, the node sends in parallel the event report over all possible trees by replicating the event report over all trees. The replicated events are propagated to the BSs and the sink. According to the events time and locations, the sink will fuse the replicated events into one single event.

V. ENERGY SAVING METHODS

The energy resources of the nodes are very limited. A critical factor in the design of the VLSSNA algorithm was to reduce the energy utilization as much as possible and to enable the network operators to select an operation mode that fits the needs and will meet the energy resources of the nodes. The following methods are used by the network. (*a*) In case that a node is able to control its transmission power, and it has more than a single candidate to become its father, it will select the most "economical" father that is not too close. (*b*) Optional Acknowledge (Ack) via "shouting" and not via the network. (*c*)Using adjustable replicated transmissions. A node can generate an Ack message that will be transmitted in parallel on one or more trees

according to its importance. (d) The ability of the initiating sensor to select randomly a tree contributes also to an even utilization of energy in the network.

VI. THE INTERACTIVE FLEXIBLE AD HOC SIMULATOR

For testing and evaluating the protocols described in this research (VLSSNA and TFA) we used the interactive Flexible Ad-Hoc Simulator (IFAS) [7]. The IFAS simulator was originally developed for evaluating the performance of ad-hoc protocols; it was adapted for sensor networks and can handle successfully thousands of sensors. In this section, we shall describe the simulator and the simulation scenarios.

Special attention was given to the following aspects: (a) enhanced visualization tools that give a full visual of the theater, zooming of selected zones, node movements and voice channels in ad hoc networks, and specific node status including queue status; (b) tracing the formation of trees; (c) tracing the events transfer and sessions in real time; (d) configuration and simulation definition via online screens; (e) definition and tracking of intruders paths and pace; (f) support of logging, debugging and analysis tools.

The enhanced visualization capabilities, unique to this simulator, contributed to the understanding of the protocols behavior, as we were able to view the progress in the field and detect unexpected behavior.

The simulator enables the user to get detailed online reports. These capabilities set afloat disruptions in specific nodes behavior as a result of their location in the field. It is possible to "kill" nodes, zoom in and out selected areas and trace explicitly certain events. Fig. 4 presents one of the trees on a terrain with 3000 sensors (the black points) placed randomly in the field.



Figure 4: Sample Tree with BS-2 as a root

The root tree is BS_2 . The intruders which are not seen in this view are crossing the field and activating the sensors. In the bottom of the screen we see the events received by the BS.

Table 2: Quantities details		
Parameter	Value	
Number of sensors	3000 sensors (except for the scalability tests)	
Field size	1Km x 1Km	
Sensor transmission range	60 meters	
Event detection range	10 meters	



Figure 5: Base stations Organization

VII. SIMULATIONS AND RESULTS

The simulation environment creates the infrastructure to analyze the following directions: (a) the efficiency and scalability of the TFA algorithm. (b) The performance of the network in the following aspects: The contribution of the aggregation process to reduce the number of messages and the delay created by nodes as a result of the aggregation process. The tests were performed using the parameters presented in Tab. 2.

Fig. 5 presents the BSs organization. BS_1 , BS_2 and BS_3 act as the trees roots and are able to broadcast "shout" messages that cover the whole terrain. In addition, the events are received via these BSs.

The localization algorithm requires using a minimum of three synchronized beacons that broadcast periodically a time event. The best organization for the beacons is in an equilateral triangle. To show an applicative possibility, we combined the functionality of two beacons with BS_1 and BS_2 and created a Support Basestation (SBS) that participates only on the localization process. The distance between BSs, BS_1 and BS_2 is *d*. Note that it is possible to use the SBS as a replacement for BS_3 .

Scalability and connectivity

The basic requirement from the TFA algorithm is to connect all sensors in the field into one single network. This set of tests is comprised of two groups:

1) Scalability tests. We run the TFA algorithm on a field with a minimal population of 1000 sensors and a maximal population of 10,000 nodes. The trees creation process succeeded to fuse all sensors into a single tree without any measurable impact on the performance.

2) Connectivity tests. The connectivity tests included two groups of tests – the first group verified that all nodes dispersed in the tree are merged into a single tree. The second group of tests checked what happens if sub-tree nodes within the tree die. In this case, the tests show that the nods that belong to the subtree of the faulty node discover the fault, and declare themselves as standalone trees. This declaration initiated the fusion process that results in a new fully connected field.

Average nodal delay

Intuitively, two factors contribute to the efficiency of the aggregation process – the nodal delay time and the number

of the tree levels that participate in the aggregation process. Fig. 6 presents the impact of the nodal delay time on the aggregation process. In this test, we measured the percentage of aggregated messages out of all event messages triggered by the intruder as a function of the nodal delay time. In this test, every inbound message received by a sensor on its way to the sink is delayed for a fixed period before it is outbounded to the next tree level. We expect the number of aggregated messages to grow as the internal delay grows. As presented, the number of aggregated messages grows significantly to 33% when the delay grows to 700ms. A small increase to 40% is achieved when the delay grows to 1200ms. Additional delay time greater than 1200ms does not contribute to the performance. Note that the increase in the performance costs a significant delay in the arrival of the alarm message to the sink.

VIII. CONCLUSIONS

The tree based connection between sensors presents a very efficient and practical way to connect between very large numbers of sensors in a sensor network. The method presented in this paper depicts also a replication tree

REFERENCES

- [1] G. Pottie and W. Kaiser, "Wireless integrated network sensors," Communications of the ACM, vol. 43, May 2000, pp. 51-58.
- [2] J. Kahn, R. Katz, and K. Pister, "Mobile networking for smart dust," In preceedings of The Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom), 1999, pp. 271-278.
- [3] J. N. Al-Karaki and A. E. Kamak, "Routing Techniques in Wireless Sensor Networks: A Survey," IEEE Wireless Communications Vol. 11 Issue 6, 2004, pp. 6-28.
- [4] Y. Ben-Asher, M. Feldman, S. Feldman, and P. Gurfil. IFAS: Interactive flexible ad hoc simulator. Simulation Modeling Practice and Theory, 15(7), 2007, pp. 817–830.
- [5] Y. Fan, A. Chen, L. Songwu, and Z. Lixia, "A scalable solution to minimum cost forwarding in large sensor networks". Proceedings of Tenth International Conference on Computer Communications and Networks, 2001, pp. 304-309.
- [6] C. Intanagonwiwat, R Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks", Proceedings of the ACM MobiCom, Boston, Ma, 2000, pp. 56-67.
- [7] J. Kulik, W. R. Heinzelman, and H. Balakrishnan, "Negotiationbased protocols for disseminating information in wireless sensor networks", ACM/IEEE Int. Conf. on Mobile Computing and Networking, Seattle, WA, Aug. 1999, pp. 174-175.
- [8] D. Braginsky and D. Estrin", "Rumor Routing Algorithm for Sensor Networks", in proceedings of the First Workshop on Sensor and Applications (WSNA), Atlanta, GA, Oct. 2002, pp. 1-12.
- [9] J. Kulik, W. Rabiner, and Hari Balakrishnan", "Negotiation-Based Protocols for Disseminating Information in Wireless Sensor Networks", Wireless Networks, vol. 8, Num. 2-3, 2002, pp. 169-185.
- [10] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy Efficient Communication Protocol for Wireless Mi-crosensor Networks," Proceedings of the Hawaii International Conference on System Sciences (HICSS '00), Jan. 2000, pp 323-327.
- [11] L. Subramanian and R. H. Katz, "An Architecture for Building Self Configurable Systems", in the Proceedings of the IEEE/ACM

mechanism that increases the redundancy of the network and ensures a very high level of connectivity.

The aggregation algorithm is targeted to reduce the number of events that are transferred from the network to the sinks. The main purpose of this algorithm is to save transmission energy. A major consideration in the decision of the quality of this algorithm is the usage of extra energy required in implementing more sophisticated algorithms, the extra requirements from the processing unit and the memory size of the sensor. In addition, transmitting more data that is required to improve the algorithm increases significantly the energy consumption. The savings should be balanced against the cost of transferring all raw events to the sinks.



Figure 6: Average Aggregation percentage Vs. Nodal Delay

Workshop on Mobile Ad Hoc Networking and Computing, Boston, MA, August 2000, pp. 63-73.

- [12] Q. Fang, F. Zhao, and L. Guibas, "Lightweight Sensing and Communication Protocols for Target Enumeration and Aggregation", Proceedings of the ACM international symposium on Mobile ad hoc networking and computing (MobiHoc), 2003, pp. 165-176.
- [13] F. Ye, H. Luo, J. Cheng, and S. Lu, L. Zhang, "A Two-tier data dissemination model for large-scale wireless sensor networks", proceedings of the annual ACM/IEEE MobiCom, 2002, pp 148-159.
- [14] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed Energy Conservation for Ad-hoc Routing," In Proceedings of the Seventh Annual ACM/IEEE International Conference on Mobile Computing and Networking, 2001, pp. 70-84.
- [15] A. Savvides, C-C Han, and M. Srivastava, "Dynamic fine-grained localization in Ad-Hoc networks of sensors," Proceedings of the Seventh ACM Annual International Conference on Mobile Computing and Networking (MobiCom), July 2001, pp. 166-179.
- [16] N. Bulusu, J. Heidemann, and D. Estrin, "GPS-less low cost outdoor localization for very small devices", IEEE Personal Communications Magazine, vol. 7. no. 5, Apr. 2000, pp 28-34.
- [17] S. Capkun, M. Hamdi, and J. Hubaux,"GPS-free positioning in mobile ad-hoc networks", Proceedings of the 34th Annual Hawaii International Conference on System Sciences, 2001 pp. 3481-3490.
- [18] R. Kumar, R. Hosam, and C. Guohong, A. Farooq, Y. Aylin and T. La Porta, "Conestion Aware Routing in Sensor Networks", Technical Report, <u>http://nsrc.cse.psu.edu/tech_report/NAS-TR-0036-2006.pdf</u>, 2006.
- [19] F. Y. Alvin Chen, S. Lu, and L. Zhang, "A scalable Solution to Minimum, Cost Forwarding in Large Sensor Networks," in Proceedings of the 10th international conference on Computer Communications and Networks, , 2001, pp 304-309.
- [20] B. Liu and D. Towsley, " A Study of the Coverage of Large-scale Sensor Networks," in IEEE International conference on Mobile Ad-Hoc and Sensor Systems, 2004, pp. 475-483.