Intelligent data aggregation in sensor networks using artificial neural-networks algorithms

A. Kulakov[*] and D. Davcev[**]

[*]Computer Science Department, Faculty of Electrical Engineering,
Skopje, Macedonia, kulak@etf.ukim.edu.mk
[**] Computer Science Department, Faculty of Electrical Engineering,
Skopje, Macedonia, etfdav@etf.ukim.edu.mk

ABSTRACT

Some of the algorithms developed within the artificial neural-networks tradition can be easily adopted to wireless sensor network platforms and will meet the requirements for sensor networks like: simple parallel distributed computation, distributed storage, data robustness and auto-classification of sensor readings. As a result of the dimensionality reduction obtained simply from the outputs of the neural-networks clustering algorithms, lower communication costs and energy savings can also be obtained.

In the paper we will propose three different kinds of architectures for incorporating the ART and FuzzyART artificial neural networks into the small Smart-It units' network. We will also give some results of the classifications of real-world data obtained with a sensor network of 5 Smart-It units, each equipped with 6 different types of sensors. We will also give results from the simulations where we have purposefully made one of the input sensors malfunctioning, giving zero or random signal, in order to show the data robustness of our approach.

Keywords: computational methods and numerics, smart mems and sensor systems, neural networks, data robustness

1 INTRODUCTION

Sensor networks systems are highly data-driven and are deployed to observe, analyze and understand the physical world. A fully centralized data collection strategy is infeasible given the energy constraints on sensor node communication, and is also inefficient given that sensor data has significant redundancy both in time, space and sensor modalities. In cases when the application demands compressed summaries of large spatio-temporal sensor data and similarity queries, such as detecting correlations and finding similar patterns, the use of a neural-network algorithm seems a reasonable choice.

The adopted neural-networks algorithms use simple computations, do not represent big burden to memory and the proposed modified ART models can be easily parameterized according to user needs for greater or lower level of details of the sensor data. Up to date, the only application of neural-networks algorithms for data processing in the field of sensor networks is the work of Catterall et al. [3] where they have slightly modified the Kohonen Self Organizing Maps model. Even this application was presented to a different kind of audience at a conference for Artificial Life. This has additionally motivated us to bring closer the work done in the field of Artificial Neural Networks for over 40 years to the community of researchers working in the field of sensor networks, since some of the problems for the processing of the sensory input data are similar.

Artificial Neural Networks that use unsupervised learning typically perform dimensionality reduction or pattern clustering. They are able to discover both regularities and irregularities in the redundant input data by iterative process of adjusting weights of interconnections between a large numbers of simple computational units (called artificial neurons).

An adopted neural network algorithm can be implemented in the tiny platform of Smart-It units, which are kind of sensor nodes or motes. Thus instead of reporting the raw-data, each Smart-It unit can send only the cluster (category) number where the current sensory input pattern has been classified. In that way a huge dimensionality reduction can be achieved depending on the number of sensor inputs in each unit (in our case it's a 6-to-1 reduction). Since the communication is the biggest consumer of the energy in the units, this leads to bigger energy savings as well.

2 ART AND FUZZYART MODELS

Several models of unsupervised Artificial Neural Networks have been proposed like Multi-layer Perceptron, Self-Organizing Maps (SOMs), and Adaptive Resonance Theory (ART) [6], [7]. Out of these we have chosen the ART models for implementation in the field of sensor networks because they do not constrain the number of different categories in which the input data will be clustered, and they do not require separate cycles of learning and classifying. Having two separate cycles is inconvenient in the presence of potentially unlimited stream of input data with no reliable method of choosing the suitably representative subset for a learning cycle. ART models of unsupervised learning include ART1 [4] for binary input patterns and FuzzyART [5] for analog input patterns. They both develop stable recognition codes by self-organization in response to arbitrary sequences of input patterns. They were designed to solve the so called stability-plasticity dilemma: how to continue to learn from new events without forgetting previously learned information. ART networks model several features of biological systems such as normalization (an ability to level out variations in intensity), contrast (an ability to detect signal mixed with noise), and both short- and long-term memory to accommodate variable rates of change in the environment.

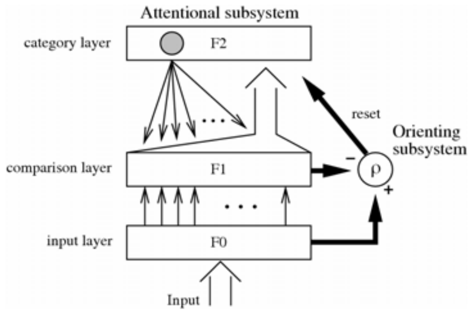

Figure 1 Architecture of the ART network.

In Figure 1 typical representation of an ART Artificial Neural Network is given. Winning F2 category nodes are selected by the attentional subsystem. Category search is controlled by the orienting subsystem. If the degree of category match at the F1 layer is lower than the so called vigilance level ρ, a reset signal will be triggered, which will deactivate the current winning F2 node thus allowing other nodes to win.

An ART network is built up of three layers: the input layer (F0), the comparison layer (F1) and the recognition layer (F2). The input layer stores the input pattern, and each neuron in the input layer is connected to its corresponding node in the comparison layer via one-to-one, non-modifiable links. Nodes in the F2 layer represent input categories. The F1 and F2 layers interact with each other through weighted bottom-up and top-down connections that are modified when the network learns. There are additional gain control signals in the network (not shown in Figure 1) that regulate its operation, but those will not be detailed here. The details about the learning process of the network can be found in [2], [4]-[7].

As a consequence of its stability-plasticity property, the network is capable of learning "on-line", i.e. refining its learned categories in response to a stream of new input patterns, as opposed to being trained "off-line" on a finite training set.

The number of developed categories can be controlled by setting the vigilance level ρ: the higher the vigilance level, the larger number of more specific categories will be created. At its extreme, if $\rho = 1$, the network will create a new category for every unique input pattern.

The strengths of the ART models include their unique ability to solve a stability-plasticity dilemma, extremely short training times in the fast-learning mode, and an incrementally growing number of clusters (categories) based on the variations in the input data. The network runs entirely autonomously, it does not need any outside control, it can learn and classify at the same time, provides fast access to match results, and is designed to work with infinite stream of data. All these features make it an excellent choice for application in wireless sensor networks.

3 PROPOSED ARCHITECTURES OF SENSOR NETWORKS

Three types of network architectures are proposed. The results of the classifications of a real-world data will be given later for each of the architectures.

3.1 One Clusterhead collecting all sensor data

First we have made this architecture to compare the work of [2] (which used Kohonen SOMs), in order to show that ART model can be used straightforwardly instead of SOMs. This model brings advantages in that we do not have to fix in advance the number of clusters (categories) that the network should learn to recognize. Here the Smart-It units send the sensory readings to one of them chosen to be a Clusterhead, where a FuzzyART network is implemented.

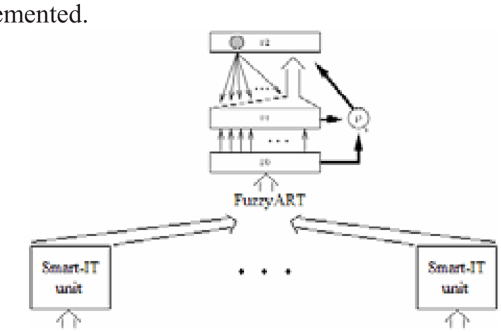

Figure 2 Clusterhead collecting all sensor data from its cluster of units

3.2 Each unit being a Clusterhead clustering data with different level of details

In this architecture each unit receives the input data from all Smart-It units in one cluster by a broadcast. Then each unit classifies the sensor data with different vigilance parameter ρ, thus providing a general overall view on the network, (smaller ρ) or more and more detailed views of the network (greater ρ). Depending on the level of details needed at the moment, the corresponding Smart-It unit can be queried depending on the level of vigilance ρ used to classify the data.

Instead of having only one Clusterhead, since the data is broadcast anyway, in this architecture all Smart-It units collect data from all over units and they all have FuzzyART implementations. So we can use different vigilance parameters ρ with which we achieve different kinds of views over the same data (coarser with smaller number of categories or more detailed with bigger number of categories).

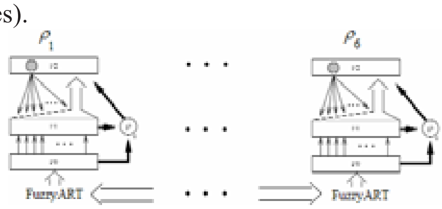

Figure 3 Redundant Clusterheads collect data at different levels of details

3.3 Clusterhead collecting only clustering outputs from the other units.

Each Smart-It unit has FuzzyART implementations classifying only its sensor readings. One of the Smart-It units can be chosen to be a Clusterhead collecting and classifying only the classifications obtained at other units. Since the clusters at each unit can be represented with integer values, the neural-network implementation at the Clusterhead is ART1 with binary inputs.

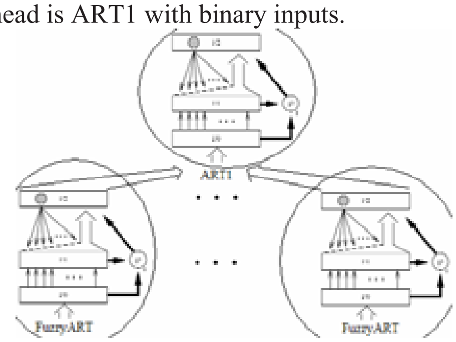

Figure 4 One Clusterhead collecting the data after they being classified at the lower level

With this architecture a great dimensionality reduction can be achieved depending on the number of sensor inputs in each unit (in our case it's a 6-to-1 reduction). In the same time communication savings benefit from the fact that the cluster number is a small binary number unlike raw sensory readings which can be several bytes long real numbers converted from the analog inputs.

Since the communication is the biggest consumer of the energy in the units, this leads to bigger energy savings as well.Another benefit from this architecture is the fact that we can view the classifications at the Clusterhead as an indication of the spatio-temporal correlations of the input data.

4 EXPERIMENTAL RESULTS

The data used in these experiments were provided courtesy of authors of [3] (the data files are available for download at: www.comp.lancs.ac.uk/~catterae/alife2002/). The details about the hardware platform can also be found in [3]. Here we will mention that the data are collected on a network of 5 Smart-It units, each equipped with a light sensor, a microphone, 2 accelerometers, a thermometer, a pressure sensor, and an RF stack providing wireless communication only for broadcast. All results presented here were produced using datasets containing real-world data taken over 1700 samples. All the experiments were conducted with complement coding of the input vector and fast learning mode. Figure 5 shows one possible classification of these input data in an architecture presented in Figure 2 with vigilance level ρ=0.93.

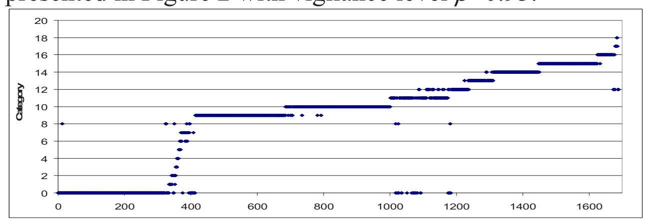

Figure 5 One possible classification of the input data

For testing the data robustness of the models, we have synthetically made one of the sensors at a time defective in way that it gives either a zero constant signal or a random measurement signal. Training was done with vigilance set to 0.93, while testing was done with vigilance set to 0.90.

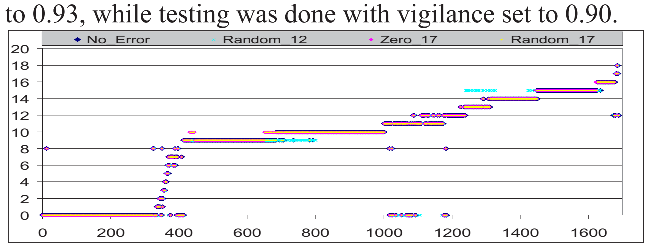

Figure 6 Different classifications when some of the sensors are defective giving zero or random values.

In Figure 6 the effects of the representative sensor errors are shown (sensor numbered 12 and sensor numbered 17, out of 30), where with the ovals are highlighted the regions where the classifications differ from the case when all sensors are functioning correctly. In Regions 1 and 2, the classification of the case when the sensor number 12 gives random values differs from the regular case, while in Region 3, the defective sensor number 17 results in different classification than the regular case. In Region 2, the cases when the 17th sensor gives random or zero values also results in different classifications.

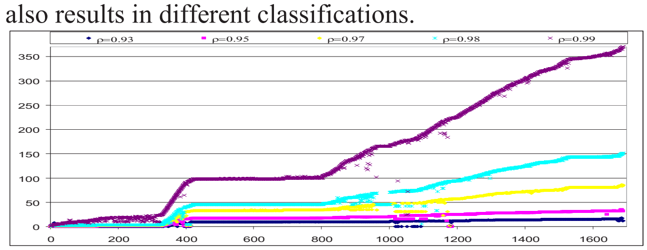

Figure 7 Redundant Clusterheads collecting data at different levels of details

Second architecture (Figure 3) provides different degrees of granularity of the input data, namely for different values of the vigilance parameter ρ (ranging from 0.93 up to 0.99 in our experiments) we get different number of output categories (from 20 up to 370) for the 1700 samples.

For the third architecture (Figure 4) we have also conducted experiments with the original data and with the synthetically made erroneous data. In Figure 8 we give the results of the classifications of the Clusterhead collecting only the classifications from the other Smart-It units. The training was done with vigilance level of 0.88, while the testing with 0.70. The results show no significant difference among the classifications when all sensors are functioning correctly or when some of the sensors give only zero or random signal (in our case sensors number 12 and 17).

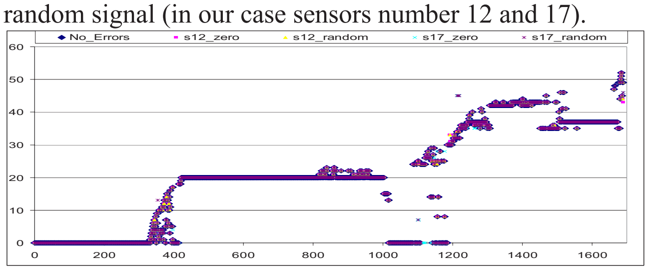

Figure 8 Results of the classifications show significant data robustness of the third architecture with one Clusterhead collecting only clustering outputs from the other units.

5 DISCUSSION

The third proposed architecture (Figure 4) with one Clusterhead collecting only clustering outputs from the other units can be generalized to a hierarchical cascade classification scheme where Smart-It units at the lowest level will be grouped in small groups of FuzzyART classifiers having one ART1 Clusterhead. Then several Clusterheads can be grouped and their outputs can be classified using a binary input ART1 classifier at a Clusterhead one-level higher and so on, up to a level where the classification will be read by a human user or stored in a database, after achieving a huge dimensionality reduction (see Figure 9).

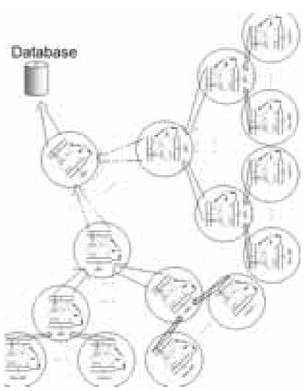

Figure 9 Hierarchical cascades of ART neural-network classifiers implemented in units of a sensor network

6 CONCLUSION

In this paper we have demonstrated a possible application of one popular model of Artificial Neural Networks algorithm (ART model) in the field of wireless sensor networks. The positive features of the ART class algorithms such as simple parallel distributed computation, distributed storage, data robustness and auto-classification of sensor readings are demonstrated within three different proposed architectures using ART1 and FuzzyART models.

One of the proposed architectures with one Clusterhead collecting only clustering outputs from the other units provides a big dimensionality reduction and in the same time additional communication saving, since only classification IDs (small binaries) are passed to the Clusterhead instead of all input samples.

Results from the simulated erroneous sensors, where we imitate defective sensors giving only zero or random output, show that the model is robust to small variations in the input.

REFERENCES

[1] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin, and S. Madden, "Distributed regression: an efficient framework for modeling sensor network data," in Proc. Information Processing in Sensor Networks, Berkeley, California, USA, 2004.

[2] G. A. Carpenter, S. Grossberg, "Integrating symbolic and neural processing in a self-organizing architecture for pattern recognition and prediction." In V. Honavar, L. Uhr (Eds.), "Artificial intelligence and neural networks: Steps towards principled prediction." San Diego: Academic Press, 1994, pp. 387-421.

[3] E. Catterall, K. Van Laerhoven, M. Strohbach, "Self-Organization in ad- hoc sensor networks: an empirical study," in Proc. of Artificial Life VIII, The 8th Int. Conf. on the Simulation and Synthesis of Living Systems, Sydney, NSW, Australia, 2002.

[4] G. A. Carpenter, S. Grossberg, "A massively parallel architecture for a self-organizing neural pattern recognition machine," Computer Vision, Graphics, and Image Processing, 1987, vol. 37, pp. 54-115.

[5] G. A. Carpenter, S. Grossberg, D. B. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," Neural Networks, 1991, vol. 4, pp. 759-771.

[6] S. Grossberg, "Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors," Biological Cybernetics, 1976, vol. 23, pp. 121-134.

[7] S. Grossberg, "Adaptive Resonance Theory" in "Encyclopedia of cognitive science," Macmillan Reference Ltd., 2000.