Distributed Pattern Recognition in Wireless Sensor Networks

Robert Richter

***Abstract*—This paper is the outcome of a literature research regarding approaches to distributed pattern recognition in Wireless Sensor Networks. The focus is to excerpt the fundamental approaches described in the papers and to give a categorization of them.**

***Index Terms*—WSN,**

I. Introduction

WIRELESS Sensor Networks (WSNs) offering a wide range of real-world applications constitute an appreciable area of research in computer science and electrical engineering. Typically, they consist of uniform electronic devices with various sensory capabilities distributed within an area of interest without any wired power supply. In [1], applications of WSNs are categorized into sampling systems and surveillance systems. The former being characterized by simple nodes gathering and forwarding uniformly sampled data to a centralized base station for aggregation and further processing, the latter being focused on event detection, tracking and classification. Energy efficiency being a central topic in WSN architectures, in general both approaches are constrained to operate with the least amount of energy consumption as possible. The area of application and system properties of surveillance systems, however, offers further possibilities of in-network processing, since the interest is in events rather than raw data. With such an approach reducing the amount of multi-hop forwarding of most times non-event containing data alongside the way to a central processing unit, which is a highly energy consumptioning process in general, is thought to be feasible.

The purpose of this paper is to summarize some recent approaches dealing with the task of node-collaborative event detection, which can be generalized to be the task of distributed pattern recognition (DPR) in WSNs.

II. Motivation and Background

Pattern recognition is a well established area within computer science. Figure 1 shows the functions generally involved in the recognition process. In the following, this general scheme is described as an adoption to electronic digital systems like WSNs usually are.

Real world data is recorded by sensors and converted to digital signals by analog-digital converters. The signals are post processed, i.e. filtered to reduce noise and/or normalized to allow generalized further processing (one should mention, that these tasks may also be performed before analog-digital conversion by analog circuits). The boundaries between feature extraction and reduction are fluent. Basically several metrics are extracted from the post processed signals that as clearly as possible define the patterns to be recognized to form a feature vector. Since not everything is useful for classification or may even complicate it, the feature vectors are reduced to the optimal amount of needed information. To give some examples, features of a signal may be amplitude, spectrum, mean, variance etc. Finally a classification takes place to map a sensed event, described by the feature vector, to a recognized class by a classifier.

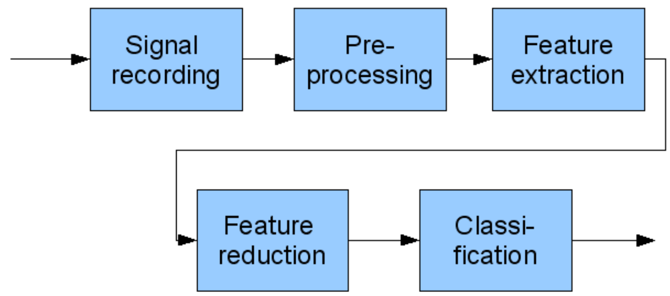

Fig. 1. General pattern recognition process

A WSN, because of its distributed structure, allows to run multiple instances of this process in parallel, but one wants to achieve a single global classification result in general. So the question of how and where to fuse this process within a WSN arises, with regards to the constrain, that for data exchange, a radio transmission has to take place, that, as already mentioned, is quite power consumptioning if related to the general power budget of a single WSN node.

In the following, some published answers to this question are presented.

III. Approaches to DPR in WSNs

A. DPR in VigilNet

In [12] Gu et al., present the DPR approach to the VigilNet project also covered in [1]. VigilNet is a WSN for tracking persons, persons with ferrous objects and vehicles primarily for military surveillance purposes in prototype phase.

VigilNets DPR-scheme comprises four hierarchy levels being sensor-level, node-level, group-level and base-level. Sensor- and node-level are implemented on each node. The group level is represented by distinct group leader nodes and the base-level is represented by a central station being the network's sink and outer world interface regarding outputs. See figure 2 for a graphical illustration.

On sensor level, three different sensor types are used being a magnetometer, a microphone and a motion sensor. The authors cover tasks of sensor imperfection compensation/calibration,

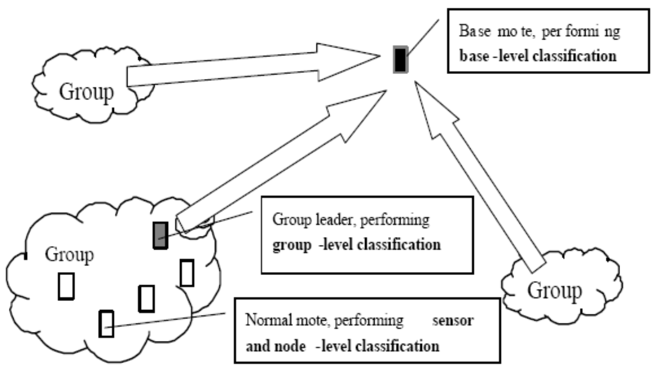

Fig. 2. VigilNet DPR approach. Adopted from [12]

i.e. temperature dependence. They post process the raw sensor data by low-order (1-tap in general) infinite impulse response filters, to raise the SNR of the signals and to compensate for system impairments. Unfortunately, the local classification procedure is not covered by the authors. However, they mention, that it involves a fusion of the three sensor outputs, to form a single classification result, one may derive, that it is threshold based at sensor level, since they deal with noise power dependent threshold adjustment to compete weather dependent power floor alteration. Also, for the motion sensor they describe their confidence vector metric to be the number of threshold oversteps within a given time frame.

Groups in VigilNet are formed by middle ware called EnvioSuite [11]. The nodes report meta information and their confidence vectors to the group leader. The most important parameter of the group-level classification is the minimum degree of aggregation (MDOA), which determines the amount of needed distinct node-level confidence vector reports (probably within a given time window). If the constraints are met, the group manager forms a confidence vector itself and sends it to the base station via multi-hop routing.

Although the approach seems to be well aligned with the real-world impairments a rolled out surveillance WSNs is confronted with, the authors somehow missed to describe their approach in the required depth concerning DPR. It is not clear, how a node level confidence vector is generated. But since this is the basic metric the group-level detection is based on, also the group-level classification mechanisms could not be extracted from the paper in detail.

B. Maximum Likelihood approaches

In [13], Duarte presents simulation results based on real-world sampled data of a classification method based on maximum joint posterior probability with distance weighting. The given evidence is based on mean and variance of spectral analysed sensor data (so Fast Fourier Transformation (FFT) is involved). Duarte gives an approximative function, the likelihood decision is based on. He sets the collaborative decision to be the weighted sum of the individual decisions. He compares the classification rate, being the correctly classified targets out of the sensed targets and the rejection rate, being the number of rejected samples (a vector of real world data) out of all samples (somehow this seems not to be equal to the more common metric of ”correct negative”) for several target distance based decisions. The used weight w for the compared approaches is set to

1) the probability of a measured evidence given target distance and SNR times the probability of target distance and SNR
2) 1 if the target distance is the lowest out of all other node’s target distance, otherwise 0
3) 1 if the target distance is lower than a predefined threshold, otherwise 0
4) 1 always

He compares this approaches by simulation and concludes that classification rates are comparable good for 1) and 2) in relation to the other approaches, but on the other hand rejection rate is to low for 1) and 2) and additionally 2) has a rather good distance error robustness.

In [5], Brooks at al. discuss some maximum-likelihood (ML) decision based DPR schemes for target classification within a tracking WSN. Their model comprises multiple nodes, each able to sense multiple ”modalities” being in fact differently post processed event signals recorded by different sensors.

Events are formed from an interval of digital time signals, where the beginning of the interval is determined by the exceeding of a threshold and the end is determined by the signal falling below the threshold. The event data is post processed (they perform a FFT) to form a complex valued ”feature vector”. Feature vectors are assumed to be Gaussian distributed, so their mean and covariance matrix are used to form the input to the ML-decision process. The output of a ML-decision is a classification of the input into a predefined set of possible targets. To focus on the interesting case of multi-node and multi-modal setup, as it can be commonly found in DPR WSN, they pose the question of how to perform the DPR task when considering global decisioning on feature vector fusion basis or on ML-decision fusion basis. For their specific prototype setup, they conclude, that intra node fusion of modals is more efficiently performed by feature vector based fusion, whereas inter node fusion is more efficiently performed by ML-decision fusion of the involved nodes. For this purpose, a managing node within a group of decisioning nodes is responsible for the combined ML-decision. They justify this conclusion by observing too much communication demand for a feature vector based inter node fusion within their setup.

C. Cooperative Fusion

In [6], Dziengel qualifies the conclusion taken within [5], with regards to the number of the feature vectors needed to be exchanged for a inter node feature vector based fusion being relatively high in the presented setup.

Dziengel considers inter node feature vector exchanges, as an

optional possibility of the network to manage uncertainties, caused by contradictory single node decisions. However, Dziengel does not perform a ML-decision making. A ML-decision making as proposed by [5] always yields the most likely classification, in the case that the taken assumptions hold (and they take the basic assumption of Gaussian processes), but is also relatively complex to perform.
Dziengels approach is less computational complex which comes with a tradeoff as a matter of principle. He proposes a fusion of local (per node) classification results, which is performed by a consensus decision.
For local classification a so called Prototype Modeller is used, which is described in [7]. The basic concept of this approach is to define a class by a prototype. A prototype is a point in multidimensional space. This space is build out of all features contained in the classification. The class' prototypes have been previously found by a k-means-algorithm based training period. An event is mapped to a class, if its distance within the multidimensional space to that class' prototype, is minimal.
If no global consensus is achievable the node responsible for the global decision requests the specific feature vectors of all nodes, the individual nodes based their local classification on. With this finer grained data base, the global decision making node repeats the classification based on a merged feature vector comprising all features from the nodes. In this case, the prototype modeller known from the local classification is reused, but is now extended to work in the grown feature space. The author names this decision making comprising classification results and feature vectors cooperative fusion. The author relates this approach to the transmitted bytes to achieve it and concludes, that this approach results in a good compromise between correct pattern recognition and transmission effort.

D. The Graph Neuron

Kahn et al. [2] propose the Graph Neuron (GN) structure as being an abstract neural network akin overlay to a WSN. The basic idea is to have a node recognize a single point in a value-position space i.e. a word $\sigma \in \Sigma^n$ with $|\Sigma| = m$ would require nm nodes to be recognized. Patterns known to a node consist of memory entries comprising its own responsibility within the value-position space but also the responsibilities of its adjacent nodes. For a given match in within a node's own responsibility, defined "ports" are listened on, using which adjacent nodes report their match result. With this information the node produces a result being "recall" if the pattern is recognized or "store" if its new. The adjacency is pattern dependent and abstractly described as edges between the nodes. Figure 3 shows a little example from the authors. A word of length three over the alphabet $\{X, O\}$ shall be recognized. The given word instance makes the corresponding nodes respond. The nodes then communicate via the ports.

For a global pattern analyses, the results of the nodes are processed by a base station.

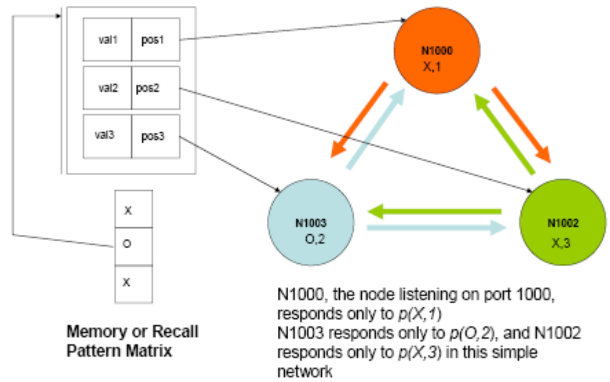

Fig. 3. An simple example of a Graph Neuron, showing the three active out of six nodes. Adopted from [2]

Kahn et al. adopt this approach to a structural health monitoring system and conclude, that their approach "has the potential to be developed into a novel application, where material objects could sense their internal states autonomously using the limited computational resources available within their embedded sensor networks, and possibly take remedial actions".
In [3] Kahn et al. propose a GN based DPR scheme to detect Distributed Denial of Service Attacks against a WSN. Their results on that may be seen as a proof of concept. However, they are gathered by simulation.
In [4], Baqer extends the described GN approach to become a Voting GN (VGN) algorithm, where also the global decision making is embedded into the network by forming committees of nodes that cooperatively solve a pattern matching problem by integrating their local results. A global result requires consensus, so the committee members negotiate until unison is achieved. This results in an exact pattern matching rather than in an approximative.
The negotiation process is described as follows: Based on the local decision, a node forms a vote vector v comprising all possible global pattern labels (it is assumed, each global pattern is allocated to a unique globally known label, to identify it), whose patterns intersect with the node's local result. This vote vector is broadcasted into the network. If a node receives such a vector, it intersects its current vote vector with the received to form an updated vote vector, which is again broadcasted. This process repeats until convergence is received such that every node's vote vector comprises the single detected global pattern label, which results in a "recall", or more than one pattern labels remain, which is interpreted as a "store". Additionally, a sleeping mode is introduced, excluding a node from negotiation process if it receives a vote vector equal to its current one, meaning it would send redundant information, which is useless.
An evaluation of this approach is made by simulation, resulting in an $O(\sqrt{n})$ shaped dependency of the needed broadcast vectors from the amount of possible patterns detectable by the network as showed in Baqers result figure 4.

The GN approach seems to embed the pattern recognition

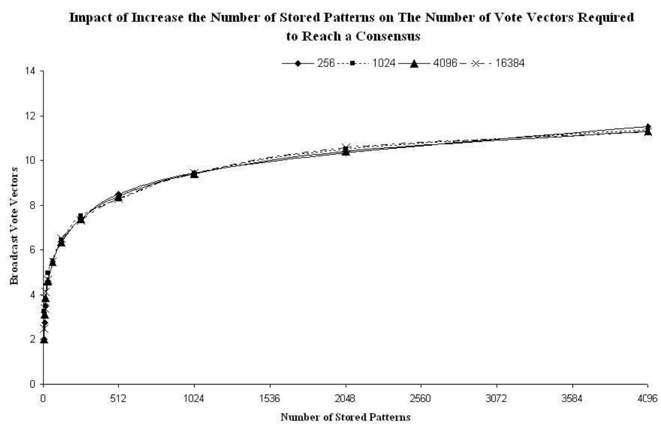

Fig. 4. An evaluation result of the VGN approach. Adopted from [4]

task nicely into the network structure, such being by design an organic knowledge-representation with neuron like processing units. However, the evaluations presented by the authors come from several assumptions, to name just fully connected nodes, perfect MAC and a priory knowledge of pattern dimension, such that a comparison to other approaches not utilizing such an elegant decision making, but therefore dealing with the real-world impairments a WSN is confronted with, can hardly be made.

E. Association-rule mining

In [8] and [9] Roemer presents an approach to discover frequent event patterns and their spatial and temporal properties using WSNs. His approach differs fundamentally from the other given approaches in terms of it's purpose is not recognition of an a priori known pattern, but the mining of unknown events and relating them to time, space, context and appearance to form a vaguely defined pattern. Therefore, a data mining technique known as association-rule mining is used in a distributed manner. It allows to constrain the reporting of in-network gathered events to the user defined parameters. A discovery node polls sensor notes within its neighbourhood for their events. The replies contain the event type and the necessary meta information for the event type, to allow the discovery note to apply the predefined association-rule. Only events passed this filtering are send to a base station. Event detection itself is assumed as given and not further considered by the authors.

IV. Comparison

A short comparison of the presented approaches to DPR in WSNs is given in table I. A good metric to evaluate the quality of an approach would be to relate true positive and false negative classifications to the amount of transmitted data, as Dziengel gave in [6], or even to the amount of energy consumption. However, the database the authors gave in their papers is not sufficient for such comparisons. But their general approaches allow a rough positioning within the plane given by classificationquality and energydemand as seen in figure 5.

TABLE I
Comparison of approaches

Approach	Distributed fusion based on	Decision making	proof of concept
VigilNet DPR	classification	na	prototype
ML with distance weight	weighted classification	ML	simulation
ML based on classification	classification	ML	prototype (SensIT)
Cooperative fusion	classification and feature vector	consensus or prototype modeller	lab prototype
VGN	feature vector (position-value pair)	consensus	simulation
Association rule mining	feature vector	association rule	na

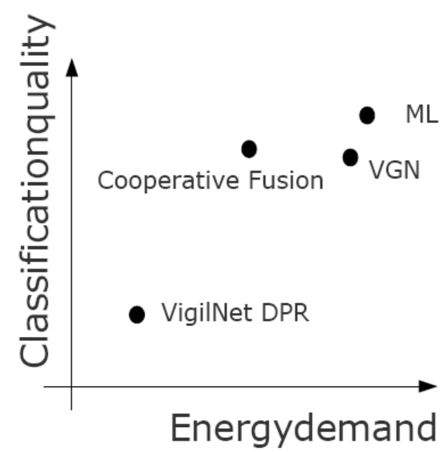

Fig. 5. Tradeoff between classificationquality and energydemand of the presented approaches to DPR

V. Conclusion and future work

Although DPR in WSNs is an obvious approach to reduce network traffic compared to sampling systems, and so, to reduce power consumption, the merger between theoretical possibilities and practical issues is still ripening. The theoretically optimal approaches seem still infeasible for real-world application due to to high processing power needs (FFT, ML-decision making), whereas the working groups dealing with realised networks have a wider view and do not focus so much to their apparently present DPR approaches or, if they do, it turns out, that several compromises have to be made.
When mentioning the wider view, a future work topic relating literature research, is to widen the view with regards to topics like group forming, target tracking, data aggregation before transmission (piggy-backing), node redundancy detection for down shutting and last but not least efficient MAC and routing protocols. All these are topics that matter when dealing with literature on DPR in WSNs and one should also put some focus on these topics, since the performance of a DPR scheme is related to them.

Mai 1, 2008

REFERENCES

[1] Tian He, Pascal Vicaire, Ting Yan, Qing Cao, Gang Zhou, Lin Gu, Liqian Luo, Radu Stoleru, John A. Stankovic, Tarek F. Abdelzaher *Achieving Long-Term Surveillance in VigilNet* University of Virginia, to appear in TOSN, 2008.

[2] Asad Khan, Zubair Baig, Matt Swift *Multi-level Pattern Recognition in Wireless Sensor Networks for Structural Health Monitoring* Faculty of Information Technology Monash University, Frankston 3199, Victoria, Australia.

[3] Z. A. Baig, M. Baqer, and A. I. Khan *A Pattern Recognition Scheme for Distributed Denial of Service (DDoS) Attacks in Wireless Sensor Networks* ICPR 2006. 18th International Conference on Volume 3, Issue , 2006 Page(s):1050 - 1054

[4] M. Baqer, A. I. Khan *Event Detection in Wireless Sensor Networks Using a Decentralised Pattern Matching Algorithm* Faculty of Information Technology Monash University, Frankston 3199, Victoria, Australia.

[5] R. R. Brooks, P. Ramanathan, A. M. Sayeed *Distributed Target Classification and Tracking in Sensor Networks* Proceedings of the IEEE Volume 91, Issue 8, Aug. 2003 Page(s): 1163 - 1171

[6] Norman Dziengel *Verteilte Ereigniserkennung in Sensornetzen* Freie Universitt Berlin.

[7] Kalton, A., Langley, P., Wagstaff, K., Yoo, J. *Generalized Clustering, Supervised Learning, and Data Assignment.* New York, NY, USA : In: ACM Press, 2001.

[8] Kay Roemer *Distributed Mining of Spatio-Temporal Event Patterns in Sensor Networks.* EAWMS / DCOSS 2006, pp. 103-116, San Francisco, USA, June 2006

[9] Kay Roemer *Discovery of Frequent Distributed Event Patterns in Sensor Networks* EWSN 2008, Bologna, Italy, January 2008

[10] Marco F. Duarte *Collaborative Signal Processing in Distributed Wireless Sensor Networks* ECE 736

[11] L. Luo, T. F. Abdelzaher, T. He, and J. A. Stankovic *Design and comparison of lightweight group management strategies in envirosuite.* In Distributed Computing in Sensor Systems (DCOSS 05), June 2005.

[12] Lin Gu, Dong Jia, Pascal Vicaire, Ting Yan, Liqian Luo, Ajay Tirumala, Qing Cao, Tian He, John A. Stankovic, Tarek Abdelzaher, Bruce H. Krogh *Lightweight Detection and Classification for Wireless Sensor Networks in Realistic Environments* SenSys '05: Proceedings of the 3rd international conference on Embedded networked sensor systems

[13] Marco Duarte, Yu-Hen Hu *Distance Based Decision Fusion in a Distributed Wireless Sensor Network* Telecommunication Systems, Vol. 26 No. 2-4, pp. 339-350, 2004