# We are IntechOpen, the world's leading publisher of Open Access books
# Built by scientists, for scientists

## 6,900
Open access books available

## 185,000
International authors and editors

## 200M
Downloads

## 154
Countries delivered to

Our authors are among the

## TOP 1%
most cited scientists

## 12.2%
Contributors from top 500 universities

**CLARIVATE ANALYTICS**
**BOOK CITATION INDEX**
**INDEXED**

**WEB OF SCIENCE™**

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

## Interested in publishing with us?
## Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

# Building Context Aware Network of Wireless Sensors Using a Scalable Distributed Estimation Scheme for Real-time Data Manipulation

Amir Hossein Basirat and Asad I. Khan
*Monash University*
*Australia*

## 1. Introduction

Wireless sensor networks are an "exciting emerging domain of deeply networked systems of low-power wireless motes with a tiny amount of CPU and memory and large federated networks for high-resolution sensing of the environment" (Welsh et al., 2004). The capability to support plethora of new diverse applications has placed Wireless Sensor Network technology at threshold of an era of significant potential growth. The technology is advancing rapidly under the push of the new technological developments and the pull of vast and diverse potential applications. The near ubiquity of the internet coupled with recent engineering achievements, are opening the door to a new generation of low-cost and powerful sensor devices which are capable of delivering high-grade spatial and temporal resolution. In that regard, distributed estimation and tracking is one of the most fundamental collaborative information processing challenges in wireless sensor networks (WSNs). Moreover, estimation issues in wireless networks with packet-loss are gaining lion share of attention over the last few years. However, due to the inherent limitations of WSNs in terms of power and computational resources, deploying any distributed estimation technique within WSNs requires modifying the existing methods to address those limitations effectively. In fact, the problem with current approaches lies in the significant increase in the computational expenses of the deployed methods as the result of increase in the size of the network. This increase puts a heavy practical burden on deployment of those algorithms for resource-constrained wireless sensor networks.

Current decentralized Kalman filtering involves state estimation using a set of local Kalman filters that communicate with all other nodes. The information flow is all-to-all with approximate communication complexity of $O(n^2)$ which is not scalable for WSNs. In this chapter an attempt is made to explore new ways in provisioning distributed estimation in WSNs by introducing a light-weight distributed pattern recognition scheme which provides single-cycle learning and entails a large number of loosely coupled parallel operations. In fact, the focus of our approach would be on a novel scalable and distributed filtering scheme in which each node only communicates messages with its neighbours on a network to

minimize the communication overhead to a high degree. In order to achieve higher level of simplicity to increase effectiveness of algorithm deployment in terms of utilizing limited available resources, a novel approach toward real-time distributed estimation is introduced, using a simplified graph-based method called Distributed Hierarchical Graph Neuron (DHGN). The proposed approach not only enjoys from conserving the limited power resources of resource-constrained sensor nodes, but also can be scaled effectively to address scalability issues which are of primary concern in wireless sensor networks. The proposed scheme is not intending to replace existing complex methods which are treated in the literature, but rather makes an attempt to reduce the computational expenses involved in the processing of the gathered in-situ information from sensor nodes. Strength of DHGN lies in the processing of non-uniform data patterns as it implements a finely distributable framework at the smallest (atomic) logical sub-pattern level. The results are easily obtained by summation at the overall pattern level. Hence, the algorithm is able to provide some sort of divide and distribute filtering process throughout the network in a fine-grained manner for minimizing the energy use. DHGN is also a highly scalable algorithm that allows real-time in-network data manipulation; essential feature for real-time data processing in WSNs. The outline for this chapter is as follows. Section 2 provides an overview of WSN technology and its current research trends. Section 3 provides a discussion on the limitations of existing approaches for data fusion and distributed estimation within WSNs. We will then introduce our proposed distributed event detection and pattern classification scheme in Section 4. Section 5 deals with performance metrics for our novel algorithm. Section 6 entails further discussion on our proposed scheme and future direction of this research. Finally, section 7 concludes the chapter.

## 2. WSN Overview

A wireless sensor network (WSN) in its simplest form can be defined as (Chong & Kumar, 2003; Akyildiz, Su, Sankarasubramaniam & Cayirci, 2002; Culler, Estrin & Srivastava, 2004) a network of (possibly low-size and low-complex) devices denoted as *nodes* that can sense the environment and communicate the information gathered from the monitored field (e.g., an area or volume) through wireless links; the data is forwarded, possibly via multiple hops relaying to a *sink* (sometimes denoted as *controller* or *monitor* ) that can use it locally or is connected to other networks (e.g., the Internet) through a gateway. The nodes can be stationary or moving. They can be aware of their location or not. They can be homogeneous or not. A traditional single-sink WSN is illustrated in Figure 1. Almost all scientific papers in the literature deal with such a definition. This single-sink scenario suffers from the lack of scalability: by increasing the number of nodes the amount of data gathered by the sink increases and once its capacity is reached the network size can not be augmented. Moreover, for reasons related to medium access control (MAC) and routing aspects, network performance cannot be considered independent from the network size. A more general scenario includes multiple sinks in the network (see Figure 2). Given a level of node density, a larger number of sinks will decrease the probability of isolated clusters of nodes that cannot deliver their data owing to unfortunate signal propagation conditions.
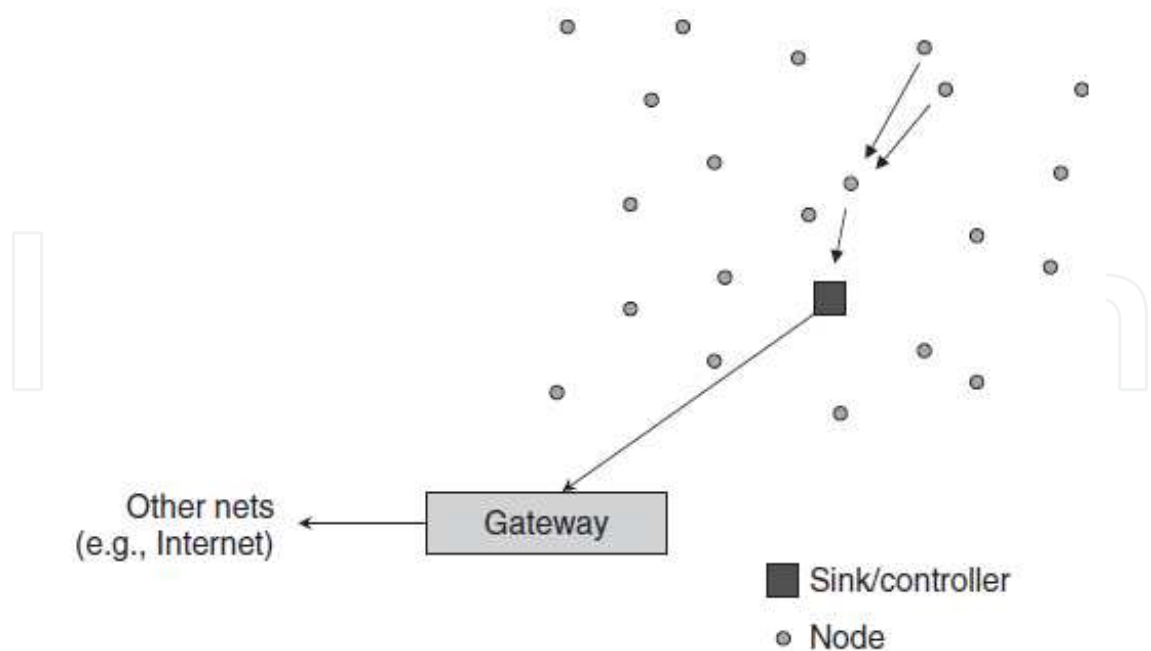
Fig. 1. Traditional single-sink WSN

In principle, a multiple-sink WSN can be scalable (i.e., the same performance can be achieved even by increasing the number of nodes), while this is clearly not true for a single-sink network. However, a multi-sink WSN does not represent a trivial extension of a single-sink case for the network engineer. There might be mainly two different cases: (1) all sinks are connected through a separate network (either wired or wireless), or (2) the sinks are disconnected. In the former case, a node needs to forward the data collected to any element in the set of sinks. From the protocol viewpoint, this means that a selection can be done based on a suitable criterion (e.g., minimum delay, maximum throughput, minimum number of hops, etc.). The presence of multiple sinks in this case ensures better network performance with respect to the single-sink case (assuming the same number of nodes is deployed over the same area), but the communication protocols must be more complex and should be designed according to suitable criteria. In the second case, when the sinks are not connected, the presence of multiple sinks tends to partition the monitored field into smaller areas; however from the communication protocols viewpoint no significant changes must be included, apart from simple sink discovery mechanisms. Clearly, the most general and interesting case (because of the better potential performance) is the first one, with the sinks connected through any type of mesh network, or via direct links with a common gateway. Both the single-sink and multiple-sink networks introduced above do not include the presence of actuators, that is, devices able to manipulate the environment rather than observe it. WSANs are composed of both sensing nodes and actuators (see Figure 3). Once more, the inclusion of actuators does not represent a simple extension of a WSN from the communication protocol viewpoint. In fact the information flow must be reversed in this case: the protocols should be able to manage many-to-one communications when sensors provide data, and one-to-many flows when the actuators need to be addressed, or even one-to-one links if a specific actuator has to be reached.
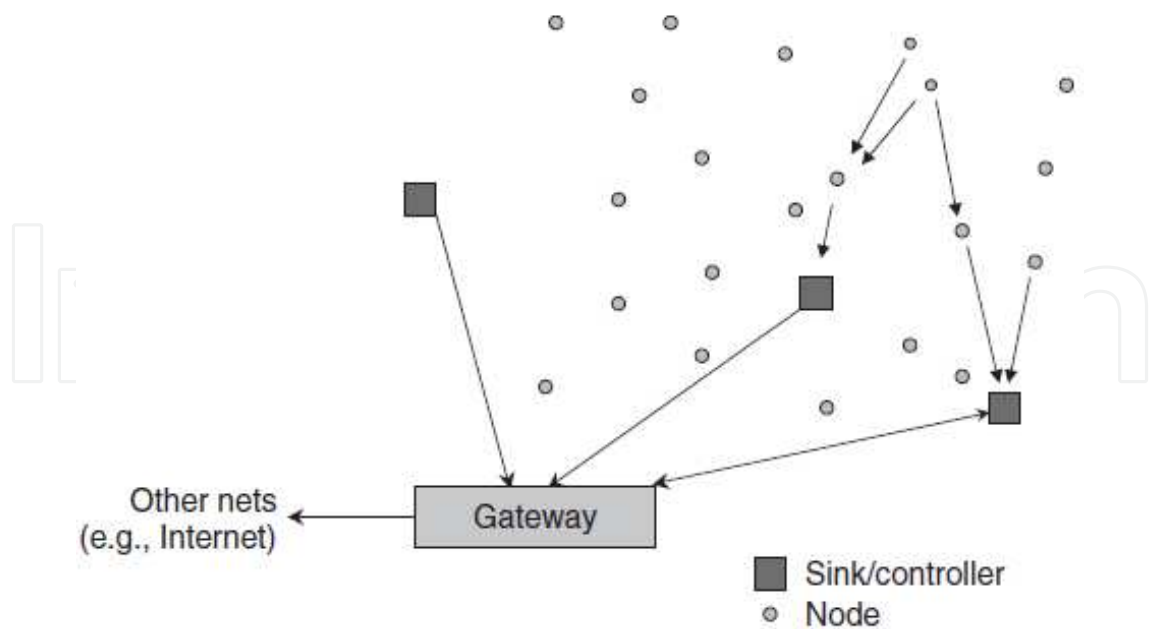
Fig. 2. Multi-sink WSN

The complexity of the protocols in this case is even larger. Given the very large number of nodes that can constitute a WSAN (more than hundreds sometimes), it is clear that MAC and the network layer are very relevant parts of the protocol stack. Tens of proposals specifically designed for WSANs have been made in the past few years. The communication protocols of a WSAN should also allow an easy deployment of nodes; the network must be able to self-organize and self-heal when some local failures are encountered.
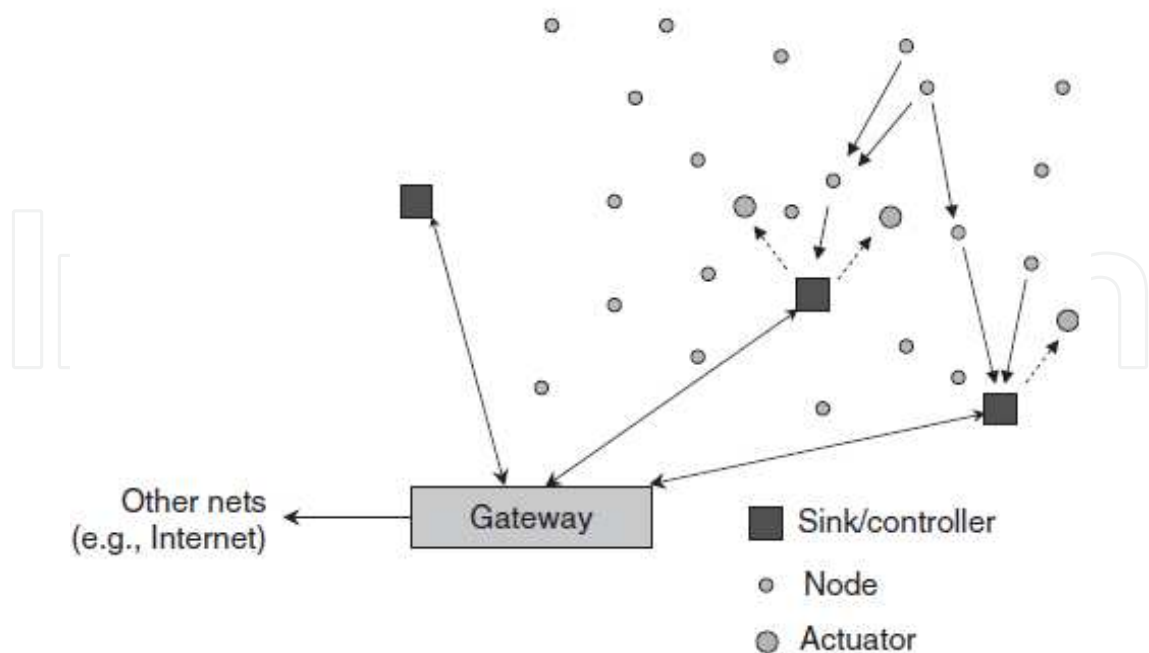


Fig. 3. Typical WSAN

## 2.1 Current and future research on WSANs

Many technical topics of WSANs are still considered by research as the current solutions are known to be non-optimized, or too much constrained. From the physical layer viewpoint, standardization is a key issue for success of WSAN markets. Currently the basic options for building HW/SW platforms for WSANs are Bluetooth, IEEE 802.15.4 and 802.15.4a. At least, most commercially available platforms use these three standards for the air interface. For low data rate applications (250 Kbits on the air), IEEE 802.15.4 seems to be the most flexible technology currently available. Clearly, the need to have low-complexity and low-cost devices does not push research in the direction of advanced transmission techniques. MAC and network layer have attracted a lot of attention in the past years and still deserve investigation. In particular, combined approaches that jointly consider MAC and routing seem to be very successful.

Topology creation, control and maintenance are very hot topics. Especially with IEEE 802.15.4, which allows creation of several types of topologies (stars, mesh, trees, cluster-trees), these issues play a very significant role. Transport protocols are needed for WSANs depending on the specific type of application. However, some of the most relevant issues investigated by research in WSANs are cross layer, dealing with vertical functionalities: security, localization, time synchronization. Basically, the research in the field of WSANs started very recently with respect to other areas of the wireless communication society, as broadcasting or cellular networks. The first IEEE papers on WSANs were published after the turn of the Millennium. The first European projects on WSANs were financed after year 2001. In the US the research on WSANs was boosted a few years before. Many theoretical issues still need a lot of investments.

The e-Mobility technology platform gathers all major players in the area of wireless and mobile communications. A strategic research agenda was released and updated in 2006. According to their views, by the year 2020 mobile and wireless communications will play a central role in all aspects of European citizen's lives, not just telephony, and will be a major influence on Europe's economy, wirelessly enabling every conceivable business endeavour and personal lifestyle. The aim of research in the field can be summarized as follows: The improvement of the individual's quality of life, achieved through the availability of an environment for instant provision and access to meaningful, multi-sensory information and content.

"Environment" means that the users will strongly interact with the environment that surrounds them, for example by using devices for personal use, or by having the location as a basis for many of the services to be used. This implies a totally different structure for the networks. Also, the context recognized by the system and it acting dynamically on the information is a major enabler for intelligent applications and services. This also means that sensor networks and radio frequency identifications (RFIDs) are increasingly important. "Multi-sensory" is related to all the users, devices, and also to the fact that the environment will be capable of sensing the users presence. Also, virtual presence may be considered, implying more sensory information being communicated, and an ideal of a rich communication close to the quality achieved in interpersonal communications or direct communications with another environment; this could also include non-invasive and context-aware communication characterizing polite human interactions. Therefore, this stretches mobile and wireless communications beyond radio and computer science into new areas of science, like biology, medicine, psychology, sociology, and nano-technologies, and

also requires full cooperation with other industries not traditionally associated with communications.

Finally, the information should be multi-sensory and multi-modal, making use of all human basic senses to properly capture context, mood, state of mind, and, for example, one's state of health. Clearly, the realization of this vision of mobile and wireless communications demands multi-disciplinary research and development, crossing the boundaries of the above sciences and different industries. Also, the number of electronic sensors and RFIDs surrounding us is quickly increasing. This will increase the amount of data traffic. The future system will be complex, consisting of a multitude of service and network types ranging across wireless sensor networks, personal area, local area, home networks, moving networks to wide area networks. Therefore, the e-Mobility vision emphasizes the key role played by WSANs as elements of a more complex system linking different types of access technologies.

ARTEMIS (advanced research & technology for embedded intelligence and systems) is the technology platform for embedded systems. The term 'embedded systems' describes electronic products, equipment or more complex systems, where the embedded computing devices are not visible from the outside and are generally inaccessible by the user. The sensor and actuator nodes of WSANs are embedded systems. According to the ARTEMIS strategic research agenda, intelligent functions embedded in components and devices will be a key factor in revolutionizing industrial production processes, from design to manufacturing and distribution, particularly in the traditional sectors. These technologies add intelligence to the control processes in manufacturing shop floors and improve the logistic and distribution chains, resulting in an increasing productivity in a wide range of industrial processes. The grand challenge in the area of sensors and actuators relates to the support of huge amounts of input and output data envisaged in the application contexts with minimal power requirements and fail-safe operation.

## 3. Data Fusion Techniques for WSANs

To fully exploit the potential of sensor networks, it is essential to develop energy-efficient and bandwidth-efficient signal processing algorithms that can also be implemented in a fully distributed manner. Distributed signal processing in a WSN has a communication aspect not present in the traditional centralized signal processing framework, thus it differs in several important aspects.

- Sensor measurements are collected in a distributed fashion across the network. This necessitates data sharing via inter-sensors communication. Given a low energy budget per sensor, it is unrealistic for sensors to communicate all their full-precision data samples with one another. Thus, local data compression becomes a part of the distributed signal processing design. In contrast, in a traditional signal processing framework where data is centrally collected, there is no need for distributed data compression.

- The design of optimal distributed signal processing algorithms depends on the models used to describe: the nodes connectivity, the nodes distribution, the knowledge of sensor noise distributions, the qualities of inter-sensors communication channels, and the underlying application metrics. Distributed signal processing over a wireless sensor network requires proper coordination and

planning of sensor computation as well as careful exploitation of the limited communication capability per sensor. In other words, distributed signal processing in sensor networks has communication aspects which are not present in the most of traditional signal processing frameworks.

- In a WSN, sensors may enter or leave the network dynamically, resulting in unpredictable changes in network size and topology. This can be due to failure between inter-sensors communication (propagation conditions, interference or non-available communication channels), duty cycling, drained batteries or nodes damages. This dynamism requires the necessity for distributed signal processing algorithms to be robust to the changes in network topology or size. These algorithms and protocols must also be robust to poor time synchronization across the network and to inaccurate knowledge of sensor locations.

There are many theoretical challenges such as establishing models, metrics, bounds, and algorithms for distributed multimodal sensor fusion, distributed management of sensor networks including auto-configuration, energy-efficient application-specific protocol designs, formal techniques for the study of architectures and protocols, representation of information requirements, and sensor network capabilities on a common mathematical framework that would enable efficient information filtering (Luo et al., 2006). From these aspects, it is clear that the design of sensor networks under energy, bandwidth, and application-specific constraints spans all layers of the protocol stacks and it is very important to have a common framework enabling all these points be taken into account even if with different approximations degrees.

In this view an example of cross-layer methodology of WSN design for environmental monitoring will be shown in the following with particular emphasis on the impact of distributed digital signal processing (DDSP) on the spatial process estimation error on one side and network lifetime on the other side. Depending on the process under monitoring and the goal of the WSAN, such as detection of distributed binary events and spatial process estimation, several techniques can be pursued with envisaging of centralized and distributed processing.

## 3.1 Distributed Estimation

Distributed estimation and tracking is one of the most fundamental collaborative information processing problems in wireless sensor networks (WSN). Multi-sensor fusion and tracking problems have a long history in signal processing, control theory, and robotics. Moreover, estimation issues in wireless networks with packet-loss have been the center of much attention lately. In most applications, the intelligent fusion of information from geographically-dispersed sensor nodes, commonly known as distributed data fusion, is an important issue. A related problem is the binary decentralized (or distributed) detection problem, where a large number of identical sensor nodes deployed randomly over a wide region, together with a global detector or fusion centre (FC), cooperatively undertake the task of identifying the presence or absence of a phenomenon of interest (PoI) (see Figure 4). Specifically, each node takes a local decision about the presence or absence of the PoI and sends its decision to the FC which is responsible for the final decision based on the information gathered from local sensors.
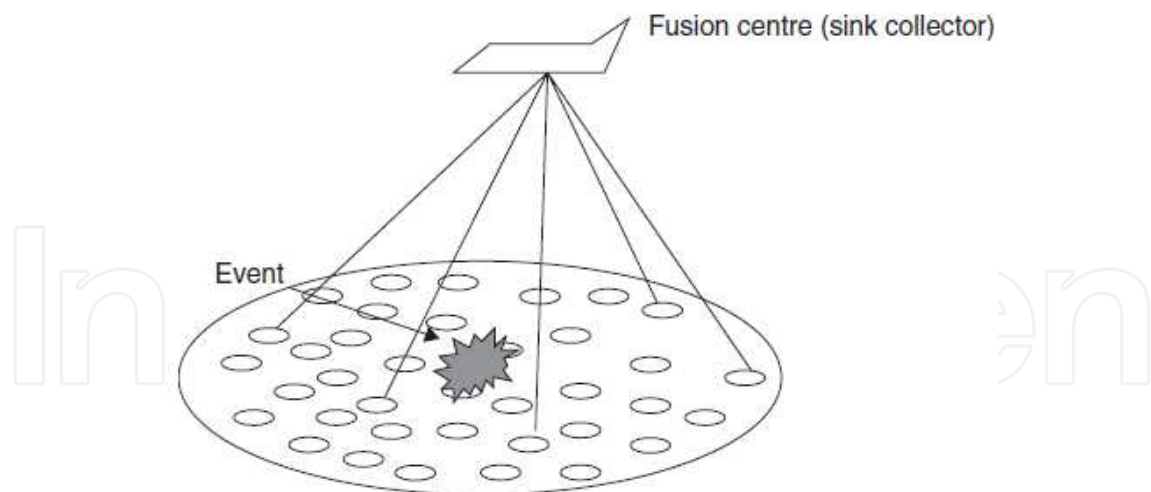
Fig. 4. Example of distributed detection scenario

Two problems have to be considered: the design of the decision rule at the FC and the design of the local sensor signal processing strategies. In the case of perfect knowledge of system parameters the design of the decision rule at the FC is a well-established task. The design of the local sensor decision rule, that is, the likelihood ratio test (LRT) threshold in binary detection, is more challenging due to the distributed nature of the system. In fact, the optimal choice of each sensor LRT threshold is coupled to each other node threshold, although nodes are not in general fully connected due to propagation effects or energy constraints. Recently it has been demonstrated that under the asymptotic regime (i.e., large number of nodes), the identical LRT threshold rule at the sensors provides the optimal error exponent if local sensor observations are independent and identically distributed (Sung, Tong & Swami, 2005). For a more complete overview of decentralized detection the reader is recommend to read Varshney, 1997; Signal Processing Magazine, Special Issue , 2006. Unlike in classical decentralized detection problems (Blum, Kassam & Poor, 1997; Varshney, 1997), greater challenges exist in a WSN setting. There are stringent power constraints for each node, and communication channels from nodes to the FC are severely bandwidth-constrained. In addition, the communication channels are no longer lossless (e.g, fading, noise and, possibly, interference are present), and the observation at each sensor node is spatially varying (Sung et al., 2005; Niu & Varshney, 2005). Recently, there has been great interest in cooperative communication (Winters, 1987; Sendonaris, Erkip & Aazhang, 2003). One may also exploit diversity associated with spatially distributed users, or simply cooperative diversity , in WSNs. In these networks, multiple sensor nodes pool their resources in a distributed manner to enhance the reliability of the transmission link. Specifi cally, in the context of decentralized detection, cooperation allows sensor nodes to exchange information and to continuously update their local decisions until consensus is reached across the nodes (Quek, Dardari & Win, 2006c; Quek, Dardari & Win, 2006b; Quek, Dardari & Win, 2006a). For example, cooperation in decentralized detection can be accomplished via the use of Parley algorithm (Swaszek & Willett, 1995). This algorithm has been shown to converge to a global decision after sufficient number of iterations when certain conditions are met. However, without a fully-connected network and given that the sensor observations are spatially varying, Parley algorithm may result in convergence to a wrong decision at most of the nodes.

### 3.2 Kalman Filtering

Although WSNs present attractive features, challenges associated with the scarcity of bandwidth and power in wireless communications have to be addressed. To perform state estimation, sensors may share these observations with each other or communicate them to a fusion center for centralized processing. In either scenario, the communication cost in terms of bandwidth and power required to convey observations is large enough to merit attention.. The Kalman filter is a very popular fusion method. Decentralized Kalman filtering (Speyer, 1979; Rao, Whyte & Sheen, 1993) involves state estimation using a set of local Kalman filters that communicate with all other nodes. The Kalman filter estimates the state x of a discrete-time controlled process that is ruled by the state-space model:

$$x[k]) = A.x[k-1] + w[k-1] \tag{1}$$

The system is influenced by process noise denoted w. The state dynamics determine the linear operator A. The state contributes to the observation y, which also includes a stochastic, additive measurement noise v:

$$y[k] = C.x[k] + v[k] \tag{2}$$

The process and measurement noises are assumed to be Normal processes with known variances W and V. The information flow is all-to-all with approxmiate communication complexity of $O(n^2)$ which can cause scalability issues for large-sclae WSNs.

## 4. Event Classification and Pattern Recognition

The reality in WSNs is that so far tremendous potential of the WSN has only been demonstrated for humble applications such as meter reading in buildings and basic forms of ecological monitoring. Reaping the full potential of this technology requires a second level of algorithms, which at present is missing. Current techniques solve the more immediate problem of conveying sensory data to a central entity known as the base station for most of the processing. This approach is not scalable and has potential to cast suspicions of big brother analysis. The WSN must therefore localize the computations within its monitored regions and communicate the computed results to small/roaming devices directly. Doing so will alleviate the scalability issue by localizing the computations and address the privacy and reliability concern through direct communications occurring within a decentralized model. Algorithms, such as the ones in this chapter, would help integrate vast networks of sensors into intelligent macro-scopes for observing our surroundings. These will bring unprecedented capabilities within our reach that transform the way we deal with phenomena occurring over large distances and inaccessible regions.

In order to achieve the broad objectives stated above, an initial step will be to develop a level of computability within the WSN whereby sensor readings can be instantly translated into event patterns and then rapidly analysed by the network (locally). This approach will entail two-fold benefit. On one hand it will enhance event detection e.g. surveillance, and target location capabilities and on the other hand it will aid in development of advanced threat detection systems for WSN, e.g. ones that can function like our biological immune system. The challenge is to evolve an approach, which can successfully detect complex real

life patterns arising from heterogeneous data sets (generated by different types of sensors) in real-time. In this regard, events of interest need to be correlated to specific pattern classes of our definition. Furthermore the asymptotic limits of the approach need to be carefully examined. For instance how far will our scheme scale with increases in the complexity of the sensory patterns and the size of the network? What will be the trade off between the scheme's sensitivity and false alarms? Can the scheme degrade gracefully with disruptions to network nodes and wireless links?

## 4.1 Graph Neuron

DHGN extends the functionalities and capabilities of Graph Neuron (GN) (Khan & Mihailescu, 2004) and Hierarchical Graph Neuron (HGN) (Khan & Nasution, 2008) algorithms. Figure 5 depicts a GN array, which is capable of converting the spatial/temporal patterns into a simple graph-based representation, in which input patterns are compared with the edges in the graph for memorization or recall operations.
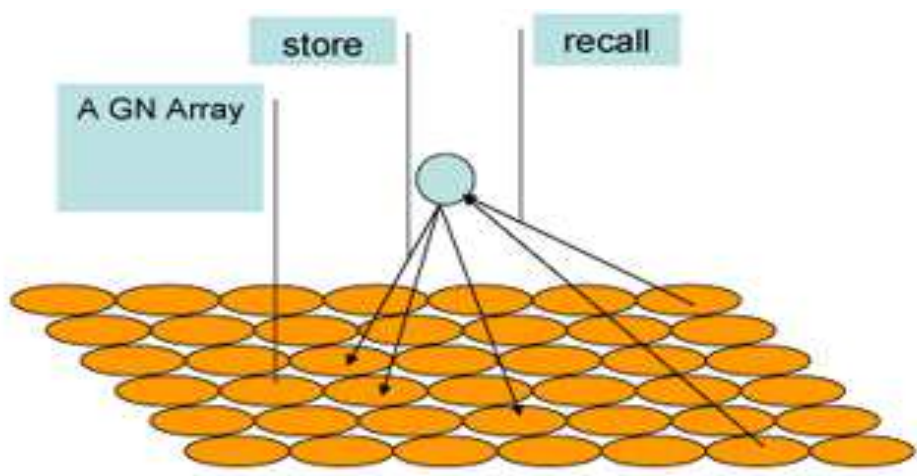


Fig. 5. Store and recall within a GN

GN is an associative memory algorithm, which implements a scalable AM device through its parallel in-network processing framework (Khan & Isreb, 2004). GN has been tested in pattern recognition applications within different types of distributed environments (Khan, Muhammad Amin, 2008a). GN makes use of graph-based model for pattern learning and recognition. One of the peculiarities of this technique is the employment of parallel in-network processing capabilities to address scalability issues effectively, which are of primary concern in distributed approaches. Graph neuron may be conceived as a directed graph in which processing nodes of the GN array are mapped to vertex set V of the graph, and links between nodes are mapped to the set of edges of the graph, E.

The communication is limited to the neighbouring nodes to decrease the amount of overhead involved and as a result, any significant increase in the number of nodes will not increase the communication overhead significantly. The information in the pattern space can be presented to each of the nodes in the graph in the form of (position, value) pair. Thus, GN array is capable of converting the spatial/temporal patterns into a simple graph-based representation in which input patterns are compared with the edges in the graph for memorization or possible recall operations (See Figure 6). Among various applications of

the GN-based AM, an input pattern in GN pattern recognition may represent bit elements of an image (Khan & Muhammad Amin, 2007) or a stimulus/signal spike produced within a network intrusion detection application (Khan, Baig & Baqer, 2006).
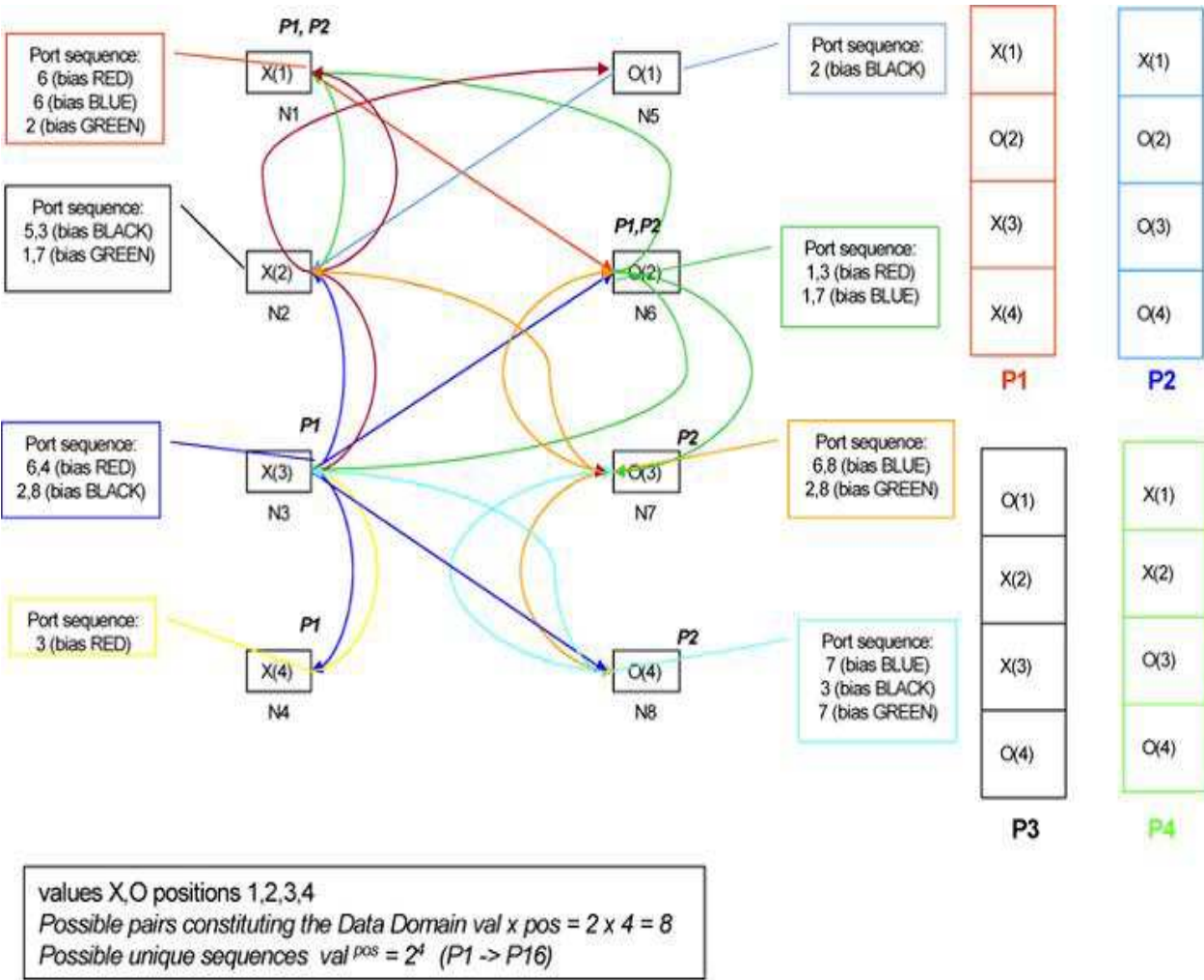


Fig. 6. An eight node GN network is in the process of storing patterns: P1 (RED), P2 (BLUE), P3 (BLACK), and P4 (GREEN)

## 4.2 Hierarchical Graph Neuron

In order to solve the issue of the crosstalk in GN model due to the limited perspective of GNs, the capabilities of perceiving GN neighbors in each GN was expanded in a new model called Hierarchical Graph Neuron (HGN) to prevent pattern interference (Khan & Nasution, 2008). Figure 7 illustrates a HGN scheme for pattern recognition with a pattern of size 7 and two possible values within the pattern being stored in the network.
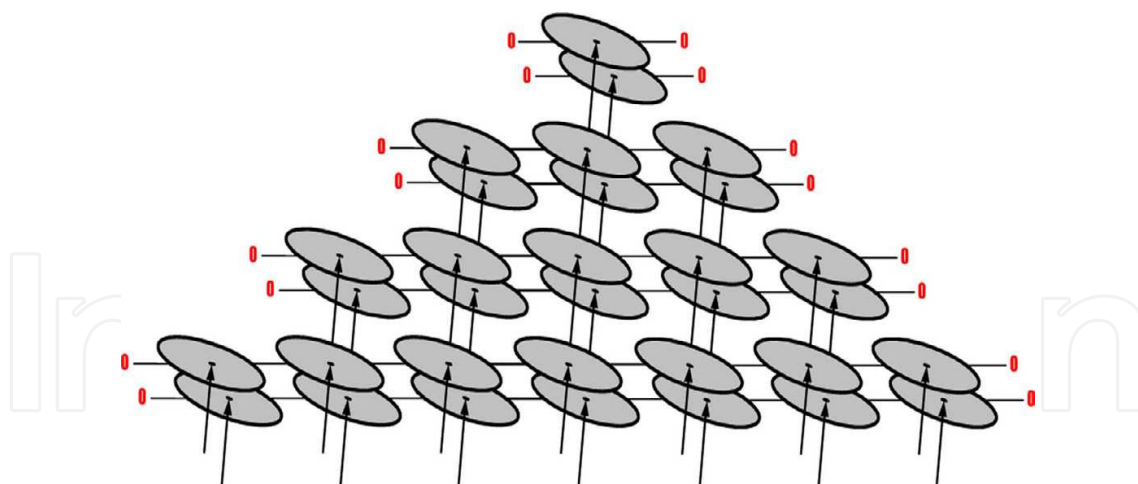
Fig. 7. Hierarchical GNs

The communication between HGN nodes is accomplished through different stages with the following fashion:

- Stage I. Each GN in the base layer will receive input pairs (value, command) sequentially. In this regard, all GNs are pre-programmed to receive these values. It should be mentioned that all the GNs in the same column will receive the same input pair, and all the GNs in the composition will receive the same command, whether store or recall.

- Stage II. Upon receiving the input, all active GNs will send a (column, row) pair to all the GNs on preceding and succeeding columns. It should be noted that, those GNs with matched IDs will issue the response; however other GNs enter the passive state, while still receiving reports from adjacent columns.

- Stage III. At this stage, all the active GNs at the edges have received one report message, while other active GNs would have received two reports from their adjacent columns. Then, each GN can determine the bias entry using input values, and report messages. Here, the GN can decide whether to store input in memory or declare it as a recall based on the command value.

- Stage IV. At this stage, after updating the bias array, each active GN will send its (row, bias index, command) value to all the GNs in the same column in the upper layer.

Finally, all the active GNs will return the (level, column, bias index) value back as the response. The process is continued by the next layer above the base layer. After repetitive execution of stages 2 to 4 by the remaining layers, the process will be stopped at one layer which is below the top layer. In general, the process consists of (S – 3) / 2 number of iterations for the last three stages, in which S denotes the size of the pattern space.


### 4.3 Distributed Hierarchical Graph Neuron

We then extended the HGN by dividing and distributing the recognition processes over the network. This distributed scheme minimizes the number of processing nodes by reducing the number of levels within the HGN. Figure 8 depicts the divide-and-distribute transformation from a monolithic HGN composition (top) to a DHGN configuration (bottom) for processing the same 35-bit patterns.
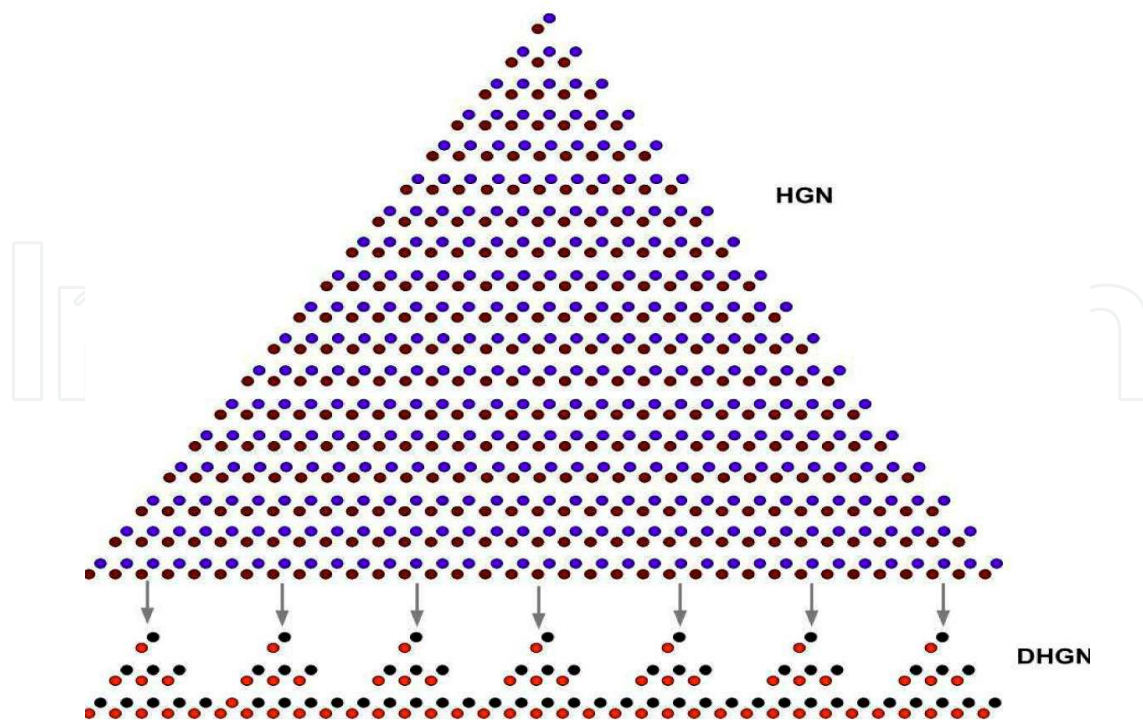
Fig. 8. Transformation of HGN structure (top) into an equivalent DHGN structure (bottom)

The base of the HGN structure in Figure 8 represents the size of the pattern. Note that the base of HGN structure is equivalent to the cumulative base of all the DHGN subnets/clusters. This transformation of HGN into equivalent DHGN composition allows on the average 80% reduction in the number of processing nodes required for the recognition process. Therefore, DHGN is able to substantially reduce the computational resource requirement for pattern recognition process – from 648 processing nodes to 126 for the case shown in Figure 8.

DHGN is in fact a single-cycle learning associative memory (AM) algorithm for pattern recognition. DHGN employs the collaborative-comparison learning approach (Khan & Muhammad Amin, 2009) in pattern recognition. It lowers the complexity of recognition processes by reducing the number of processing nodes (Khan & Muhammad Amin, 2007). In addition, pattern recognition using DHGN algorithm is improved through a two-level recognition process, which applies recognition at sub-pattern level and then recognition at the overall pattern level (See Figure 9). DHGN is already implemented in a grid environment (Khan & Muhammad Amin, 2008). We plan to adopt DHGN algorithm for our novel data estimation model in WSNs. DHGN allows the recognition process to be conducted in a smaller sub-pattern domain, hence minimizing the number of processing nodes which in turn reduces the complexity of pattern analysis. In addition, the recognition process performed using DHGN algorithm is unique in a way that each subnet is only responsible for memorizing a portion of the pattern (rather than the entire pattern). A collection of these subnets is able to form a distributed memory structure for the entire pattern. This feature enables recognition to be performed in parallel and independently. The decoupled nature of the sub-domains (subnets) is the key feature that brings dynamic scalability to data estimation within WSNs.
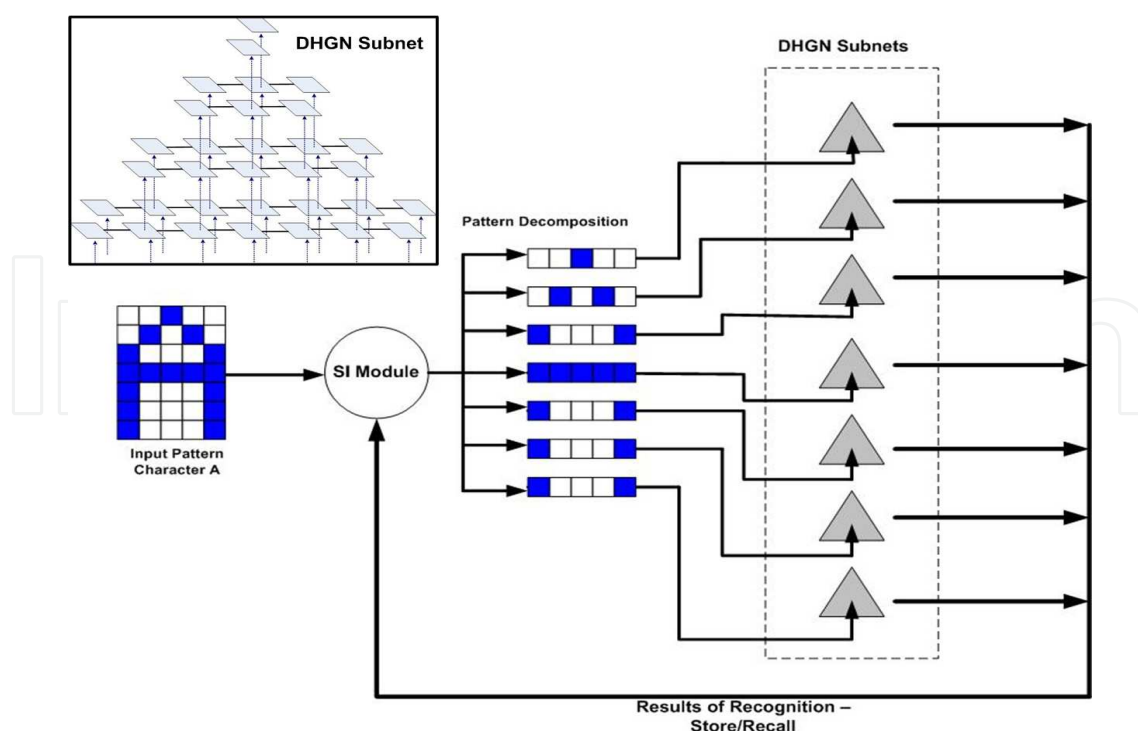
Fig. 9. DHGN distributed pattern recognition architecture

In fact, we envisage a vast associative memory network of wireless sensors to support user defined search and classification algorithms. Data in WSN may be viewed as uniform or non-uniform in nature. A uniform data block will incur the same computational cost as all the others. However in many instances, at least one data subset might be required to be known with a better quality of approximation or greater detail than the others, e.g. it may be invoked in the query process more often than the others. In this case the cost associated with data subset will be different. Larger variation can lead to inefficient use of the distributed resources. Strength of DHGN lies in the processing of non-uniform data patterns as it implements a finely distributable framework at the smallest (atomic) logical sub-pattern level. The results are easily obtained by summation at the overall pattern level. DHGN is also a highly scalable algorithm that incorporates content addressable memory within a clustered framework. Hence algorithmic strengths of the current estimation approaches can be investigated for the first time in combination with DHGN's single-cycle learning mechanism. This one shot approach will allow real-time in-network data manipulation; essential feature for data processing in WSNs.

## 5. Performance Metrics

By redesigning data processing models, data records are treated as patterns which enable data storage and retrieval by association over and above the existing simple data referential mechanisms. Hence, processing the input data and handling the dynamic load is handled by using a distributed pattern recognition approach that is implemented through the integration of loosely-coupled computational networks, followed by a divide-and-distribute approach that allows distribution of these networks within WSNs dynamically.

### 5.1 Scalable Simulation of the DHGN Application in WSNs:

Several tests are performed to ensure the accuracy of DHGN algorithm while evaluating its effectiveness in terms of complexity, timing and fault rate. It is essential that these software programs are fully integrated and tested as proof of concept that our approach will indeed work when applied to a dynamic WSN. In order to achieve this goal, we will formulate a wireless sensor network environment for executing our algorithms over very large numbers of GN nodes. Doing so will allow us to determine the asymptotical limits of our approach, find any problems that only arise when the algorithms are scaled-up, and test for robustness and scalability in face of dynamic changes to WSN configuration.

Our scheme relies on communications among adjacent nodes. The decentralized content location schemes will be implemented for discovering adjacent nodes in minimum number of hops. A GN based algorithm for optimally distributing DHGN subnets (clusters or sub-domains) among the WSN nodes will be provided to automate the boot-strapping of the distributed application over the network. Doing so will also provide means for investigating dynamic load balancing over the network. For this matter, a database of stored patterns is constructed. Patterns can be assumed as abstract notations for various parameters in the network. The user is able to configure the length and number of elements in the pattern, and based on these parameters, the program will generate a user-specified number of random patterns. At this stage, the user can manually enter patterns to be checked against the database. The program is not only capable of determining whether the pattern is matched with the previous stored ones or not, but is also able to determine the sub-patterns of the input pattern which have been visited before. This capability provides a tool for determining the level of distortion in the input pattern, which in turn adds to the robustness of the proposed algorithm significantly.

### 5.2 One shot pattern recognition within WSNs:

In contrast with hierarchical models proposed in the literature, DHGN's pattern matching capability and the small response time, that remains insensitive to the increases in the number of stored patterns, makes this approach ideal for WSNs. Moreover, the DHGN does not require definition of rules or manual interventions by the operator for setting of thresholds to achieve the desired results, nor does it require heuristics entailing iterative operations for memorization and recall of patterns. In addition, this approach allows induction of new patterns in a fixed number of steps. Whilst doing so it exhibits a high level of scalability i.e. the performance and accuracy do not degrade as the number of stored pattern increases over time. Its pattern recognition capability remains comparable with contemporary approaches. Furthermore all computations are completed within the pre-defined number of steps and as such the approach implements one shot, i.e. single-cycle or single-pass, learning. Tests have shown that the scheme has an acceptable level of pattern recognition accuracy in this context (Khan & Muhammad Amin, 2008b). As it is clearly depicted in Figure 10, the system's response (i.e. recognition times) does not trend upwards with the increase in the number of stored patterns. In other words, performance of the algorithm was not affected by the increase in the number of memorized patterns or the WSN network size.
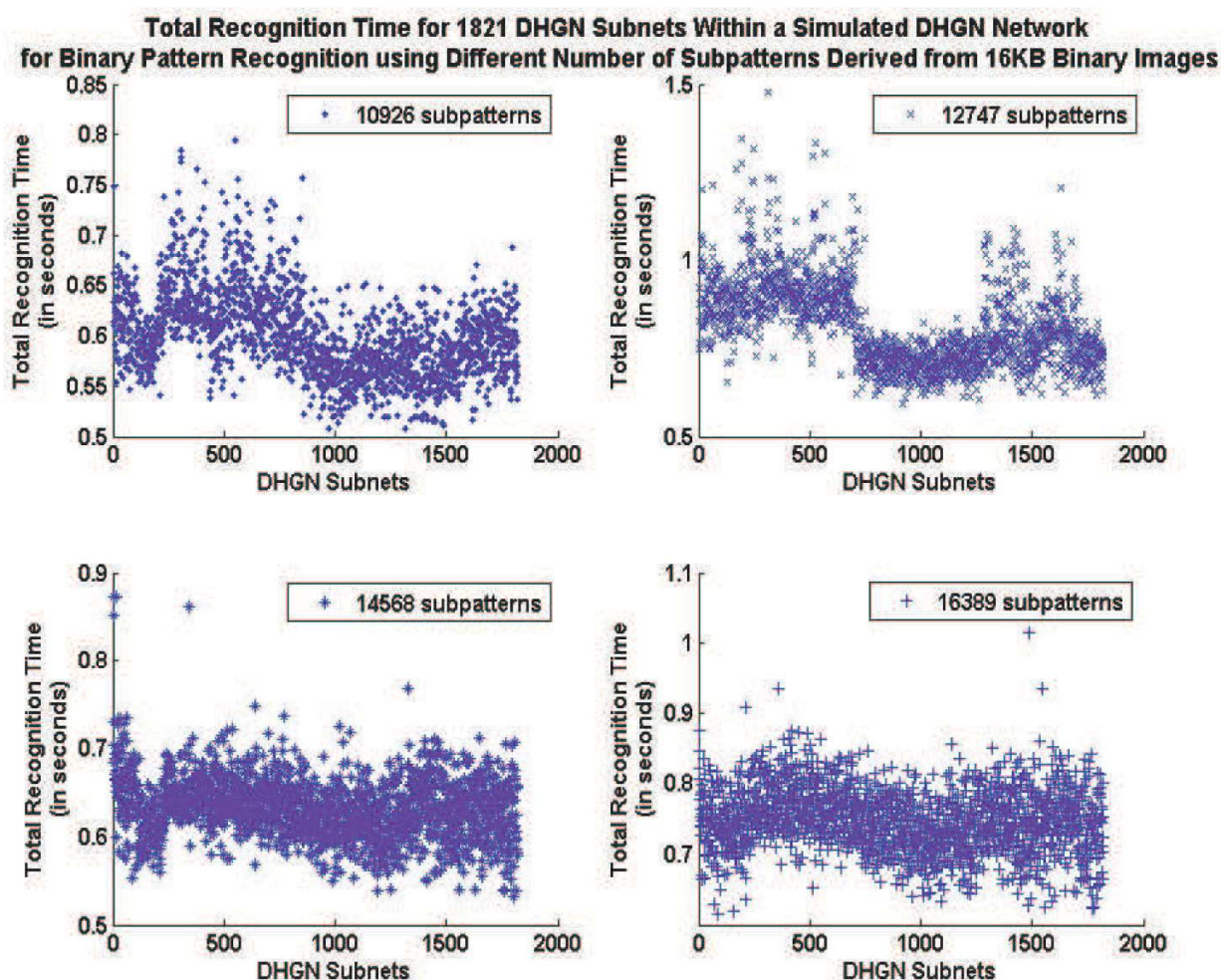
Fig. 10. Pattern recognition timings for storing a large number of 16KB images

### 5.3 One Shot Learning versus Back Propagation Network:

A comparative analysis with a Back-Propagation (BP) network layer and the input layer of the HGN highlights the ability of the GN layer to assimilate new patterns into the older patterns constructively (Khan & Nasution, 2008). The GN layer continues to improve its accuracy, as more and more patterns are stored, and exhibits better accuracy and performance as compared to the BP algorithm (See Figures 11 and 12). For performing a comparison approach, a BP perceptron layer was tested for single-cycle learning using a piecewise-linear activation function with a slope of 2.5 and a threshold of 0 for optimal results. These results were then compared with a single GN layer as shown in Figures 11 and 12 respectively. The perceptron layer and the GN layer were both trained once-only using up to 10,000 fully randomized patterns with a uniform distribution. Pattern sizes of 10 and 20 with random distortions (noise) were used in the tests. Each of these pattern sizes was associated with 16 possible values. Hence these pattern sizes corresponded to $16^{10}$ or 40 bit ($2^{40}$) and $16^{20}$ or 80 bit ($2^{80}$) patterns respectively. The accuracy was calculated as the percentage difference between the pattern size and the number of bits with recall errors.

Building Context Aware Network of Wireless Sensors Using
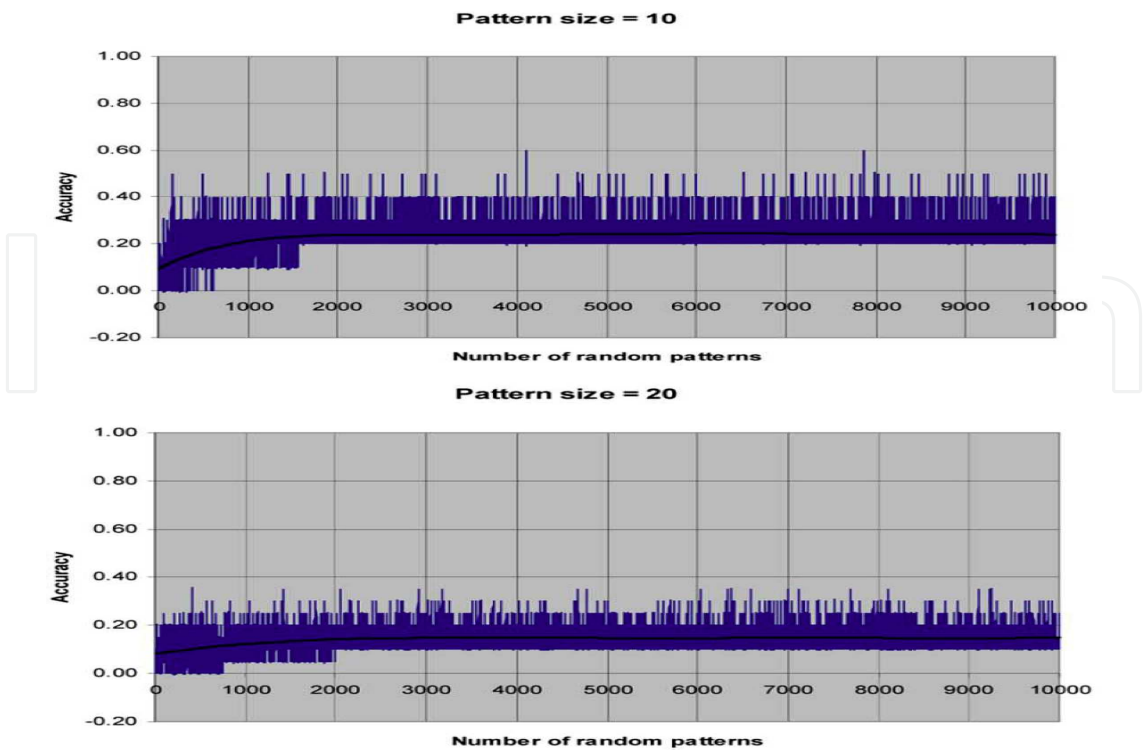a Scalable Distributed Estimation Scheme for Real-time Data Manipulation

17



Fig. 11. Recall Accuracy of BP algorithm, after being fed with up to 10,000 fully randomized patterns
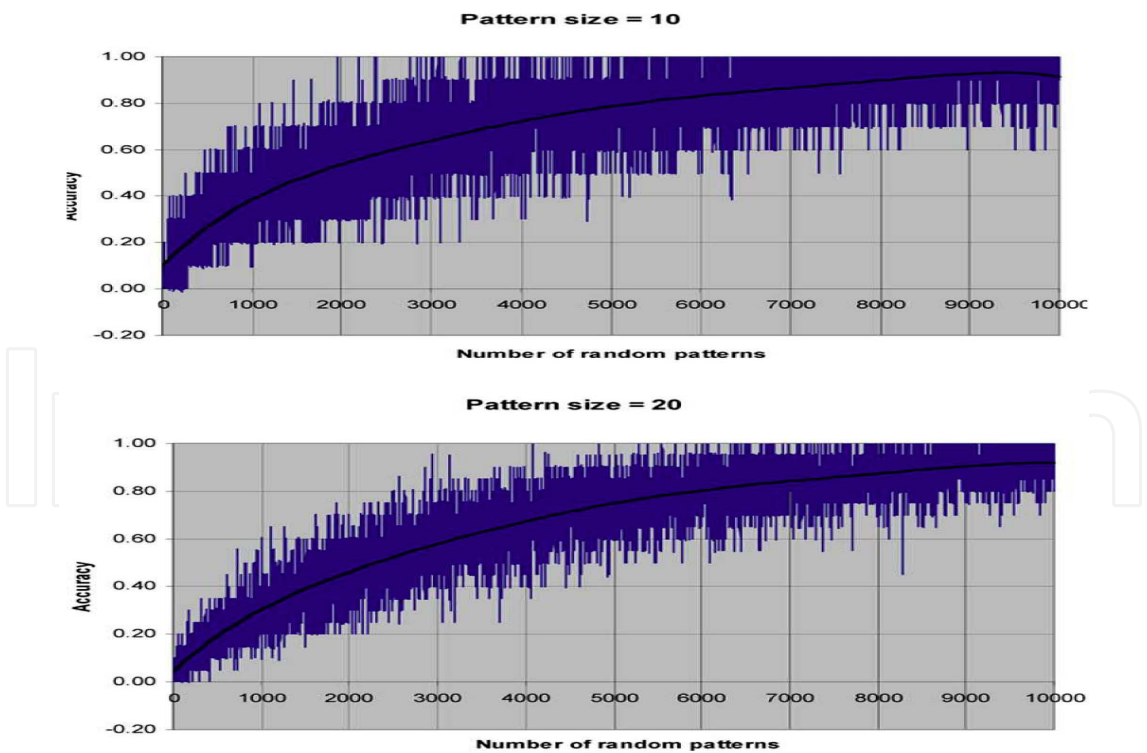


Figure 12. Recall Accuracy of single-cycle GN, after being fed with up to 10,000 fully randomized patterns (Adapted from [28])

It may be noted from Figures 11 and 12 that both the BP and the GN networks achieved comparable accuracy for the first 500 patterns. It can be seen from Figure 12 that the accuracy of GN in recognizing previously stored patterns increased as more patterns were stored (greater improvement with more one shot learning experiences). The GN in both the cases achieved above 80% accuracy after all the 10,000 patterns (with noise) had been presented. On the other hand, Figure 11 shows that the accuracy of BP network failed to improve after 2,000 patterns. Also, the recall accuracy of the perceptron layer noticeably deteriorated as the larger pattern size of 20 (80 bits) was introduced.

### 5.4 Superior Scalability:

Another important aspect of HGN is that it can remain highly scalable. In fact, its response time to store or recall operations is not affected by an increase in the size of the stored pattern database. Figure 13 illustrates the results from a HGN simulation in this regard. Figure 14 is based on the mathematical model derived in (Khan & Natusion, 2008), which corroborates the simulation results. The flat slopes in Figures 13 and 14, which show that the response times remain insensitive to the increase in stored patterns, represent the high scalability of the scheme. Hence, the issue of computational overhead increase due to the increase in the size of pattern space or number of stored patterns, as is the case in many graph-based matching algorithms will be alleviated in HGN, while the solution can be achieved within fixed number of steps of single cycle learning and recall.
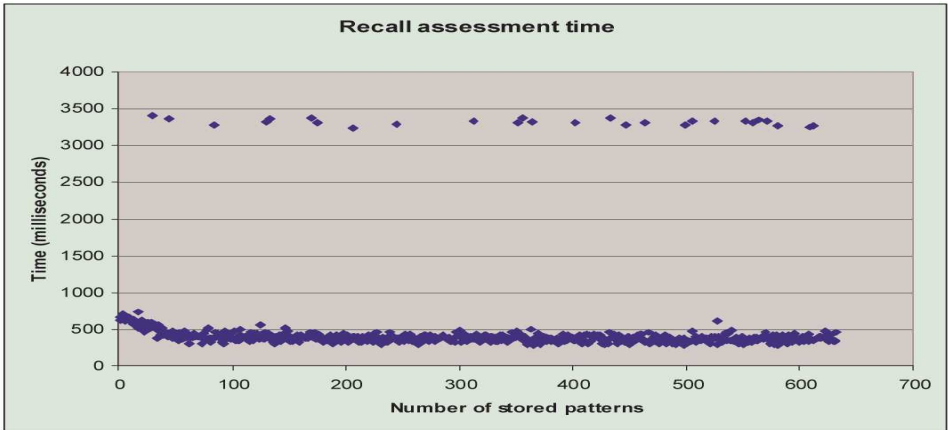


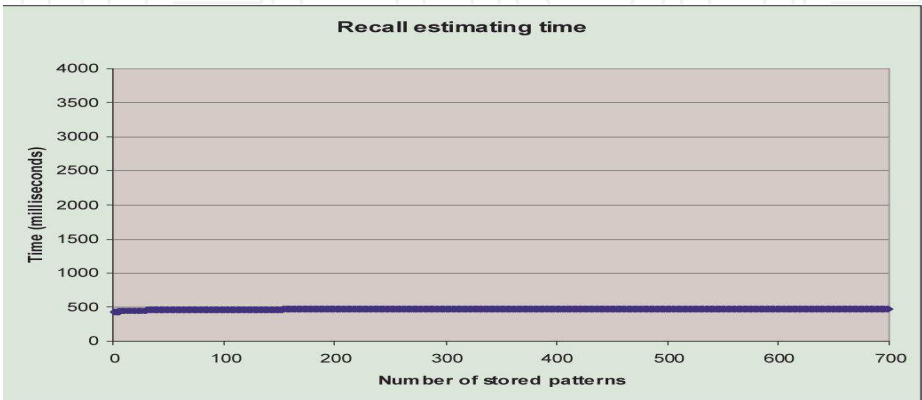Fig. 13. Actual recall time for an analysed composition



Fig. 14. Estimated recall time for a HGN composition

## 6. Discussion and Future Research

Our proposed approach is proven to work as a real-time pattern recogniser for the WSN. This is the only scheme which can operate within resource-constrained nodes of a WSN and still deliver speed and accuracies comparable to conventional pattern analysers. Furthermore our approach allows new patterns to be added at anytime without recourse to retraining. In contrast with contemporary distributed estimation approaches, our approach allows induction of new input patterns in a fixed number of steps. Whilst doing so it exhibits a high level of scalability i.e. the performance and accuracy do not degrade as the number of stored pattern increases over time. Its pattern recognition capability remains comparable with contemporary approaches. Furthermore all computations are completed within the pre-defined number of steps and as such the approach implements one shot, i.e. single-cycle or single-pass, learning. The one shot learning within this method is achieved by side-stepping the commonly used error/energy minimization and random walk approaches. The network functions as a matrix that holds all possible solutions for the problem domain. The network after applying our algorithm can find the solution in a single-cycle (i.e. fixed number of steps). Our approach finds and refines the initial solution by passing the results through a pyramidal hierarchy of similar arrays. In doing so it eliminates/resolves pattern defects, with distortions up to 20% being tolerated. Previously encountered patterns are revealed whilst new patterns are memorized without loss of stored information. In fact the pattern recognition accuracy continues to improve as the network processes more sensory inputs. HGN though highly suited for the WSN is limited by the well-known problem afflicting the hierarchical schemes i.e. large increases in network nodes as the pattern size and/or complexity increases. The distributed version of this scheme, DHGN, retains HGN's one-shot learning characteristic and reduces the computational complexity of our hierarchal scheme by distributing the recognition process into smaller clusters i.e. dividing-and-distributing simple pattern recognition processes across the network. The computational complexity of the scheme is significantly reduced and the accuracy is generally improved by using this strategy. The distributed scheme also lowers storage capacity requirements per node and incurs lesser communication cost thus improving the response-time characteristic. A comparison with the Hopfield network shows that our approach generally improves speed of recognition and accuracy. The distributed scheme also allows better control of the network resources – varying from coarse grained to very fine grained networks. This scheme compares well with contemporary approaches such as SOM and SVM in terms of speed and accuracy.

There are several benefits and advantages in our DHGN implementation for event detection and distributed estimation within WSN network. Our approach offers low memory consumption for event data storage using simple bias array representation. Furthermore, this scheme only stores sub-patterns or patterns that are related to normal event, rather than keeping the records of all occurring events. We have also shown that our approach is most effective for small sub-pattern size, since it uses only a small portion of the memory space in a typical physical sensor node in WSN network. In addition to this efficient memory usage, DHGN also eliminates the need for complex computations for event classification technique. With the adoption of single-cycle learning and adjacency comparison approaches, DHGN implements a non-iterative and lightweight computational mechanism for event recognition and classification. The results of our performance analysis have also shown that DHGN recognition time increases linearly with an increase in the number of processing elements

(PEs) within the network. This simply Lightweight Event Detection Scheme using Distributed Hierarchical Graph Neuron in Wireless Sensor Networks reveals that DHGN's computational complexity is also scalable with an increase in the size of the sub-patterns. DHGN is a distributed pattern recognition algorithm. By having this distributed characteristic, DHGN would be readily-deployable over a distributed network. With such feature, DHGN has the ability to perform as a front-end detection scheme for event detection within WSN. Through divide-and-distribute approach, complex events could be perceived as a composition of events occurring at specific time and location. Our approach eventually would be able to be used in event tracking as our proposed scheme has been demonstrated to perform efficiently within an event detection scheme such as forest fire detection using WSN. Despite all its benefits, DHGN has its own limitations. Firstly, DHGN simple data representation would requires significant advanced pre-processing at the front-end of the system. This might not be viable for strictly-resource constrained sensor nodes, where processing capability is very limited. In addition, DHGN single-hop communication for event detection scheme is not viable for large area monitoring, due to high possibility of communication error due to data packet loss during transmission. Our existing DHGN implementation has also been focusing on supervised classification. However, there is a need for unsupervised classification technique to be deployed for rapid event detection scheme. Overcoming the DHGN distributed event detection scheme limitations would be the path of our future research direction. We intend to look into event tracking scheme using DHGN distributed detection mechanism, as well as providing unsupervised classification capability for rapid and robust event detection scheme. Furthermore, we are looking forward into implementation of this scheme in large-area monitoring using multi-hop communication strategy.

## 7. Conclusion

The development of distributed estimation scheme within WSN has been made viable with the advancement in communication, computational, and sensor technologies. However, existing estimation/recognition algorithms fail to achieve optimum performance in a distributed environment, due to its tightly-coupled and computationally intensive nature. In this chapter, we have presented our readily-distributable distributed estimation scheme for WSN network which is known as Distributed Hierarchical Graph Neuron (DHGN). Throughout our studies, we discover that DHGN is able to perform recognition and classification processes with limited training data and within a one-shot learning. These DHGN features have given added-value for implementing this scheme within a lightweight distributed network such as WSN. Current implementation of DHGN in event detection using WSN has been focusing on the front-end processing, in which detection could be carried out earlier using the available wireless sensor nodes. Our approach differs from other existing estimation schemes in which major processing steps are conducted at the base station. By having a front-end detection, our proposed scheme is able to alleviate the computational costs experienced by the centralised-processing undertaken by the base station. In this chapter, we have also discussed the advantages and limitations of our proposed scheme. The future direction of this research lies in the development of a complete distributed estimation scheme that incorporates front-end detection and back-end complex event analysis. We foresee our DHGN distributed estimation scheme as a complete estimation scheme and recognition tool that is deployable over different types of event detection schemes on WSN networks.

## 8. References

Akyildiz , I. F. , Su , W. , Sankarasubramaniam & Cayirci , E. ( 2002 ). A survey on sensor networks. *IEEE Communications Magazin* , pp. 102 – 114 .

Blum , R. S. , Kassam , S. A. & Poor , H. V. ( 1997 ). Distributed detection with multiple sensors – Part II: Advanced topics. *Proceedings of IEEE*, pp. 64 – 79.

Chong , C.-Y. & Kumar, S. P. ( 2003 ). Sensor networks: evolution, opportunities, and challenges . *Proceedings of IEEE*, pp. 1247 – 1256 .

Culler , D. , Estrin , D. & Srivastava , M. ( 2004 ). Overview of sensor networks. *IEEE Computer*, pp. 41 – 49 .

Khan, A.I., Mihailescu, P. (2004), Parallel Pattern Recognition Computations within a Wireless Sensor Network, *In Proceedings of the 17th Intl. Conference on Pattern Recognition (ICPR'04)*, Cambridge, United Kingdom: IEEE Computer Society.

Khan A.I., Isreb M., Spindler R.S. (2004). A parallel distributed application of the wireless sensor network, *In Proceedings of the 7th Int'l Conf. on High Performance Computing and Grid*.

Khan, A.I., Baig Z.A., Baqer M., (2006). A pattern recognition scheme for distributed denial of service (DDOS) attacks in wireless sensor networks, *In Proceedings of the 18th International Conference on Pattern Recognition.*

Khan, A.I., Muhamad Amin, A.H. (2007). One Shot Associative Memory Method for Distorted Pattern Recognition, *In 20th Australian Joint Conference on Artificial Intelligence*, Gold Coast, Australia, pp. 705-709.

Khan, A.I., Muhamad Amin A.H., (2008a). Commodity-Grid Based Distributed Pattern Recognition Framework, *Sixth Australasian Symposium on Grid Computing and e-Research (AUSGRID 2008)*, Wollongong, NSW, Australia.

Khan, A.I., Muhamad Amin, A.H., (2008b). Single-cycle image recognition using an adaptive granularity associative memory network, *Lecture Notes in Artificial Intelligence*, vol 5360, Springer-Verlag, Berlin Germany, pp. 386-392.

Khan, A.I., Nasution, B. B. (2008) ,IEEE Transactions on Neural Networks, 212-229.

Khan, A.I., Muhamad Amin, A.H., (2009). Collaborative-Comparison Learning for Complex Event Detection using Distributed Hierarchical Graph Neuron (DHGN) Approach in Wireless Sensor Network, *In 22nd Australasian Joint Conference on Artificial Intelligence (AI'09)*, Melbourne, Australia.

Luo , Z-Q. , Gastpar , M. , Liu , J. & Swami , A. ( 2006 ). Distributed signal processing in sensor networks [from the guest editors] . I EEE Signal Processing Magazine, pp. 14 – 15.

Niu , R. , Chen , B. & Varshney , P. ( 2006 ). Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks. *IEEE Transactions on Signal Processing*, pp. 1018 – 1026 .

Quek, T. Q.-S., Dardari, D. & Win, M. Z. (2006*a* ). Cooperation in bandwidth-constrained wireless sensor networks. In *Proceedings of the IEEE International Workshop Wireless Ad Hoc Sensor Networks, New York, NY, USA* .

Quek, T. Q.-S., Dardari, D. & Win, M. Z. (2006 *b*). Energy efficiency of cooperative dense wireless sensor networks. In *Proceedings of the International Wireless Communication sand Mobile Computing Conference, IWCMC 2006, Vancouver, Canada* , pp. 1323–1330.

Quek, T. Q.-S., Dardari, D. & Win, M. Z. (2006*c*), Energy efficiency of dense wireless sensor networks: To cooperate or not to cooperate. In *Proceedings of the IEEE International Conference on Communications, ICC 2006, Istanbul, Turkey*, Vol. 10, pp. 4479–4484.

Rao, B. S. Y., Durrant-Whyte, H.F., and Sheen, J. A. (1993). A fully decentralized multi-sensor system for tracking and surveillance. *International Journal of Robotics Research*, 20–44.

Sendonaris , A. , Erkip , E. & Aazhang , B. ( 2003). User cooperation diversity – Part II: Implementation aspects and performance analysis. *IEEE Transactions on communications*, pp. 1939 – 1948.

Speyer. L. (1979). Computation and transmission requirements for a decentralized linear-quadratic-gaussian control problem. *IEEE Transactions on Automatic Control*, 266–269.

Sung , Y. , Tong , L. & Swami , A. ( 2005 ). Asymptotic locally optimal detector for large-scale sensor networks under the Poisson regime. *IEEE Transactions on Signal Processing*, pp. 2005 – 2017.

Swaszek, P. & Willett , P. ( 1995 ). Parley as an approach to distributed detection. *IEEE Transactions on Aerospace and Electronic Systems*, pp. 447 – 457.

Welsh, M. , Malan, D. , Duncan, B., Fulford-Jones, T.,  and Moulton, S. (2004). Wireless Sensor Networks for Emergency Medical Care", *GE Global Research Conference, Harvard University* and Boston University School of Medicine, Boston, MA.

Winters, J. H. ( 1987 ). On the capacity of radio communication systems with diversity in Rayleigh fading environment. *IEEE Journal Selected Areas in Communication*, pp. 871 – 878 . SAC

**Wireless Sensor Networks: Application-Centric Design**

Edited by Yen Kheng Tan

ISBN 978-953-307-321-7

Hard cover, 492 pages

**Publisher** InTech

**Published online** 14, December, 2010

**Published in print edition** December, 2010

Over the past decade, there has been a prolific increase in the research, development and commercialisation of Wireless Sensor Networks (WSNs) and their associated technologies. WSNs have found application in a vast range of different domains, scenarios and disciplines. These have included healthcare, defence and security, environmental monitoring and building/structural health monitoring. However, as a result of the broad array of pertinent applications, WSN researchers have also realised the application specificity of the domain; it is incredibly difficult, if not impossible, to find an application-independent solution to most WSN problems. Hence, research into WSNs dictates the adoption of an application-centric design process. This book is not intended to be a comprehensive review of all WSN applications and deployments to date. Instead, it is a collection of state-of-the-art research papers discussing current applications and deployment experiences, but also the communication and data processing technologies that are fundamental in further developing solutions to applications. Whilst a common foundation is retained through all chapters, this book contains a broad array of often differing interpretations, configurations and limitations of WSNs, and this highlights the diversity of this ever-changing research area. The chapters have been categorised into three distinct sections: applications and case studies, communication and networking, and information and data processing. The readership of this book is intended to be postgraduate/postdoctoral researchers and professional engineers, though some of the chapters may be of relevance to interested master's level students.

**How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Amir Basirat and Asad I. Khan (2010). Building Context Aware Network of Wireless Sensors Using a Scalable Distributed Estimation Scheme for Real-time Data Manipulation, Wireless Sensor Networks: Application-Centric Design, Yen Kheng Tan (Ed.), ISBN: 978-953-307-321-7, InTech, Available from: http://www.intechopen.com/books/wireless-sensor-networks-application-centric-design/building-context-aware-network-of-wireless-sensors-using-a-scalable-distributed-estimation-scheme-fo

# INTECH
open science | open minds