Towards Developing Attentive Wireless Sensor Networks

Renita Machado, Sirin Tekinay
Department of ECE, New Jersey Institute of Technology
{rmm23, sirin.tekinay}@njit.edu

Abstract

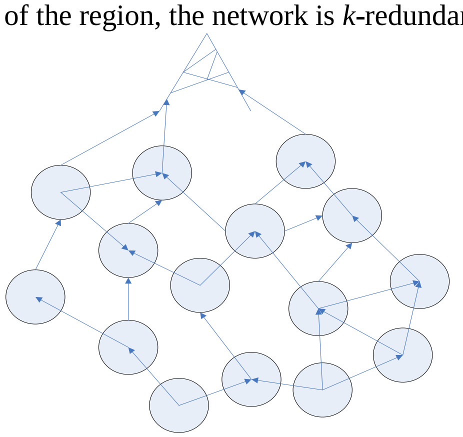
The large, rapidly growing field of wireless sensor networks (WSNs) offers the ability to collect and process massive amounts of information from various environments. This distributed data gathering and computation with the help of tiny, power-limited devices enables their use in surveillance, target detection and various other monitoring applications. In this context, the role of a sensor network can be viewed as that of a system that pays attention to a phenomenon of interest. Thus, the current body of literature on WSNs falls into two major categories: developing networks that a) pay attention to the environment to detect the phenomenon under consideration and b) improving the quality of attention paid by WSNs to these phenomena. In this paper, we summarize a theoretical framework for the context of attention in WSNs. This paper is the first step to develop a foundation for understanding the association between the nature of attention in WSNs and their real-world applications.

1. Introduction

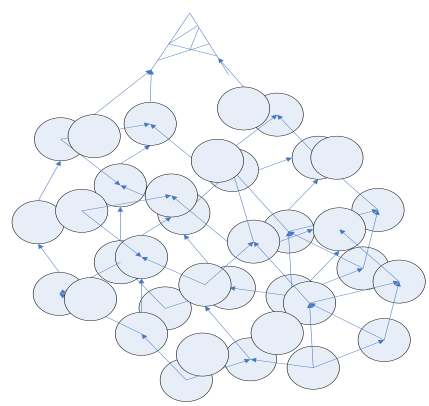
In this paper, we introduce the concept of attention in wireless sensor networks by framing a relationship between the nature of attention at the cognitive level and the parallel data-gathering and processing functions carried out by WSNs. Wireless sensor networks are aptly named for their ability to sense the deployment region, gather data and use it for higher levels of processing. Multi-hop links or a single direct link is used to route this gathered data to a central base station (sink) in order to reconstruct the desired parameters of the deployment region [1]. The power-limited nature of sensor nodes effectively constrains the processing and data dissemination that are necessary to achieve sensing objectives of reliable network operation with the objective of prolonging network lifetime. This constraint has spawned research in deployment, signal processing, communication and networking within WSNs. In particular, these issues can be classified within the context of attention into two main categories. The first category deals with developing networks that pay attention to the environment: i.e. the range of WSN applications. This is evident in the study of WSNs developed for habitat monitoring [2], weather detection [3] and structural monitoring [4] to name a few. The other category deals with improving the quality of attention paid by the WSN to the phenomenon under consideration. Deployment, density control, routing, data processing and security are themes used to improve the quality of attention paid by the WSN to the environment. Our work unifies current research in terms attention: the fundamental ability of sensor networks to pay attention and process data gathered from attentive sensing to fulfill sensing objectives. In our knowledge this is the first work that addresses WSN applications and performance as a function of attention paid by the network. Below, we provide a brief introduction to the nature and scope of WSNs and then outline the analogy between the limits of attention in human cognitive science and the limits of data gathering, processing and routing in WSNs. Wireless sensor networks (WSNs) are networks of tiny, power-limited nodes equipped with sensors, actuators and transceivers that are deployed for large-scale data gathering and processing. The nodes gather data with the help of the sensors and route this data to a central base-station or sink that collects the data from all nodes in the network for higher processing. A sample depiction of a WSN is shown in Figure 1, where the nodes are represented by their circular coverage areas. There are two major sensing scenarios: continuous sensing in which the nodes continuously gather data and route it to the sink and event-driven sensing where the nodes respond to the base station's request for data, viz. what was the temperature in region P of the deployment region at 10 am? For a WSN to be efficient, it has to satisfy the sensing objectives, viz. temperature measurement, intrusion detection, etc. in the deployment region while achieving maximum possible network lifetime.

The most commonly used measure of network lifetime is the time until the first node runs out of battery energy. The causes for a node failure arise from a combination of the following tasks a node routinely undertakes in the network: energy spent in reception of signals from other nodes or the sink, the transmission energy in transmitting the sensed data to the nearest nodes or the sink, energy expenditure in sensing the environment and in some cases, the energy spent on data processing at individual nodes which constitutes the computation expenditure. These various modes of operation can be classified as the 'awake' state in contrast to the 'sleep' state where a node can be powered off for energy conservation. Since the transmission and reception power in a wireless sensor node is much greater than the computation/sensing expenditure, an obvious design technique might be to reduce the transceiver energy expenditure or schedule nodes to enter the 'sleep' state; however this requirement might result in networks not being connected.

One important question, which is the subject of ongoing research in WSNs, concerns the optimization of the tradeoffs between network lifetime and inter-node and/or node-sink communication while maintaining coverage and connectivity throughout the deployment region. The coverage provided by a sensor node refers to the area sensed by the sensor. 100% coverage in the deployment region refers to the situation where nodes cover/sense the entire region without any sensing 'holes' (i.e. no nodes sensing certain parts of the deployment region). Redundancy of nodes in the deployment region, wherein a part of the region is covered (sensed) by more than one node is another factor that determines node coverage and connectivity requirements. Formally, when k nodes, (where $k > 1$) are deployed to sense a part of the region, the network is k-redundant (Figure 1b).



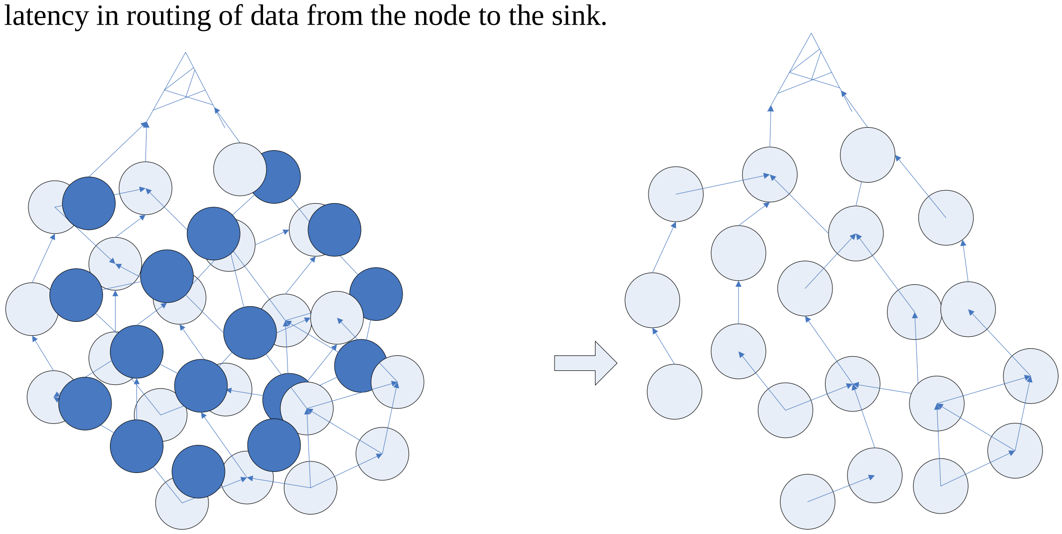
(a)



(b)

Figure 1. A wireless sensor network (WSN). (a) shows a WSN with a base station and wireless sensor nodes. The arrows indicate possible routing paths in the network. (b) shows redundancy in the WSN.

This redundancy may arise to random deployment strategies (scattering nodes from a height on remote terrains) or deterministic deployments, where dense networks of nodes are deployed in the deployment region, e.g. networks of cameras for intrusion detection. In the absence of redundancy, all nodes in the network might have to stay in the 'awake' state to achieve coverage and connectivity in the network. Redundancy presents ways to let some nodes 'sleep', while others stay 'awake' to perform data gathering, computation and routing. Figure 2a shows redundant nodes in the 'sleep' state (darkened circles). The equivalent network without the sleeping nodes is depicted in Figure 2b. This alternation between the 'sleep' and 'awake' modes' of operation helps in increasing network lifetime [5]. Connectivity implies that every node in the network is connected to at least one other node in the network. The connectivity is a function of the distance between two nodes and the randomness in the wireless channel. A *k*- connected WSN results when there can be *k* paths any two nodes *a* and *b*, such that with the removal of *k*-1 nodes, the nodes *a* and *b* are still connected. While the 'sleep' mode of operation is power-efficient, it is the 'awake' mode of operation that is responsible for data gathering, computation and routing in the WSN. The optimization problem here is to develop networks that are connected for efficient routing of data with minimum latency in routing of data from the node to the sink.



(a) (b)

Figure 2. Power management by 'sleep' scheduling. The darkened nodes indicate nodes in the 'sleep' state. The equivalent network is shown in (b).

From the above brief introduction to the structure and operation of WSNs, we see an analogy between the working of WSNs and human cognition systems in terms of the data gathering, computation and routing (Figure 3). Though the human cognitive system does not face similar constraints of working with power-limited data gathering units, it encounters constraints on the attention that can be paid to the environment. In this paper we examine the analogy between cognitive attention that deals with the visual sensory input and WSNs that gather data from the deployment region for various applications. The relevance of attention to WSNs merits a series of questions that we frame to set the tone of the rest of this paper.

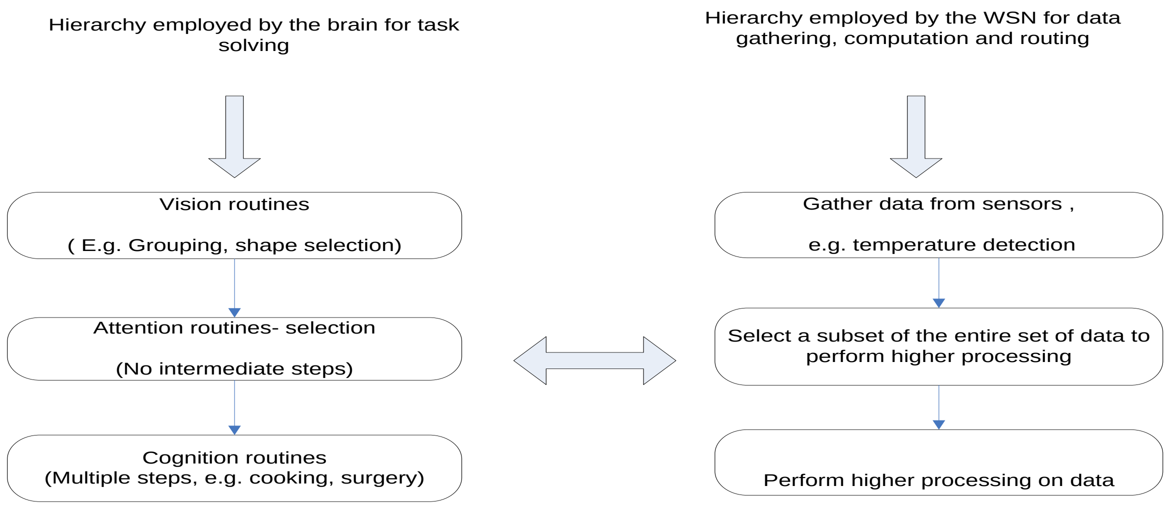


Figure 3. Analogy of task solving in human cognition [6] and the WSN

The first question that the framework evokes is: How do we extend the concept of attention to wireless sensor networks? One of the essential prerequisites for WSN applications is to create architectures for reliable data gathering and processing, so that the network fulfils sensing objectives while achieving longer network lifetimes with nodes that have constraints on battery energy and processing power. The scope of WSN applications includes data gathering in environment monitoring, surveillance, target detection and intrusion and medical applications. Though these application scenarios are unique, they all display a common feature: the WSN pays attention to the deployment region to obtain information about the parameter of interest in the sensing application. For instance, a temperature monitoring WSN pays attention to the temperature in the deployment region. In an intrusion detection system, nodes are supplied with the data set corresponding to intruder identification and the network attentively scans the environment and reports intruders to the base station. This paradigm of attention can be applied to many other sensing scenarios as well.

Having extended the concept of attention to WSNs, the next question that arises is: How do we quantify attention in WSNs? The best interpretation of this is the density of 'awake' nodes that sense the environment for any of the two main sensing scenarios: event-driven and continuous. An 'awake' node gathers data from the area covered by its sensing radius, communicates the data to the nearest node or base station.

A higher density of 'awake' nodes results in a network that is *k*-connected and depending on the deployment pattern in the region, it is also *k*-redundant. In this interpretation, attention in WSNs can be used to study efficiency of WSN operation as a function of node deployment, data processing algorithms and routing protocols. Similar to attention in humans, where a higher amount of it is linked to greater efficiency, creating attention-paying WSNs improves the network operation. This may be only trivially true, and we elaborate on this in section 3 of the paper.

Finally, how do we improve attention? It is helpful to understand the structure of attention to answer this question. The structure of attention routines defines limits to the amount of information available for higher processing and hence limits the amount or the capacity of attention [6]. In this paper, we draw on the work of [6] where the author summarized three independent limits on the information available for higher processing. These three limits are the capacity, acuity and the coding singularity of the selection region which commands attention. While the framework of attention described is not unique in that, there exist different types of attention, we focus on the attention routines that lie as an intermediate step between vision routines and cognition routines. The vision and cognition routines represent the first and last steps of a hierarchy that the brain employs to solve tasks. Within the set of attention routines, the author in [6] focused on one particular routine, selection and described it with the help of the three limits of capacity, acuity and coding singularity. We use the same approach to quantify attention limits in WSNs and detail them in section 2 of the paper.

The idea of employing cognition to wireless networks is not new; cognitive radio [7] is already being researched for wireless communication as a means to improve the utilization of scarce radio spectrum. In the tradition of WSN research, cognition can be applied to a broader framework where network applications resemble the attention paid by human sensory systems to the environment. Although this is the first formal attempt to defining cognitive WSNs, the existing research in WSNs is wide enough to be encapsulated in the framework of cognitive WSNs. The understanding of what constitutes cognitive WSNs and using the analogy between the limits of attention in human cognition and the limits of data gathering and processing in WSNs shapes the rest of this paper.

1.1. Objectives of the study on the analogy between attention in human cognition and WSNs

The study of wireless sensor networks within the context of attention is important to achieve the following three objectives:

Capacity: If the network is deployed for continuous sensing, what is the density of information that can be sensed by the network? What element of it can be used for higher processing that reliably fulfils the objectives of the sensing operation?

Acuity: In case of multiple objects in a tracking application, what is the minimum spacing between objects that can permit access and detection of the object of interest?

Coding singularity: A third and less obvious objective is to study the sensing resolution of the network. How do we focus on the features of the desired phenomenon from the entire selection region? The answer to this lies in accurate recognition of the phenomenon despite its seemingly sparse nature of description as encountered in most real-world sensing applications and developing ‘attention-paying’ WSNs.

The organization of this paper with respect to each of the above limits is as follows: Section 2 describes the capacity limit of attention. In section 3, we address the acuity and coding singularity limits and show the relationship between them. Section 4 provides emerging directions for future research and concludes the paper.

2. The capacity limit

In [6], the author showed that the capacity limit of attention in human cognition is set by the constraints of representing the initial and final routines in awareness. In WSNs, the capacity of a WSN may refer to the amount of information sensed by the network for the duration of network operation, network throughput or the transport capacity. In addition to being a function of the deployment pattern in the sensing region [8-9], we show that the capacity is also a function of the reportability. With respect to data gathering networks, the capacity is often measured in terms of the transport capacity or network throughput. In a WSN with a given density of 'awake' nodes that are sensing the environment in event-driven or continuous sensing applications, the information capacity of the network is proportional to the number of 'awake' nodes. However, the reportability required of nodes to transmit data to a sink reduces the capacity, in part due to the receiver encoding ability. This feature of the capacity limit in attention is also shown in [1]. In [1], the authors show that the amount of data received at a single receiver from the network of sensors in the deployment region depends on the density of sensors. They consider a data gathering applications, where the receiver reconstructs a snapshot of the sensed field from the data received. The compression rate at the encoder poses a constraint on the reconstruction of the sensed field, since this rate is less than the transport capacity of the network. They show that as the sensor density increases, there is more correlation in the data leading to greater compression at the encoder. However, since the single-receiver transport capacity of the receiver remains constant, the amount of time it takes to transport the sensed field/ reconstruct a snapshot of the field does not decrease, but goes to infinity.

The other factor contributing to reduction in capacity is the requirement to accurately reproduce the spatial and temporal nature of the sensed environment from the data gathered. Without this constraint, the data obtained from nodes would be compressed at the sink after an amount of time dictated by the density of nodes and efficiency of the compression algorithms used, surpassing even the encoding limit at the receiver.

3. The effects of crowding

In this section, we show the effect of crowding on attention paid by WSNs with the help of the acuity and the coding singularity limits.

3.1. Acuity

Acuity in human cognition refers to the limit imposed by crowding on selecting an object from a region of interest. In WSNs, acuity has been studied in terms of visual acuity of networks of camera sensors [10]. However, acuity in the context of an attention limit can have further implications. Acuity is an important issue for detecting/tracking applications in WSNs. In the case of a multiple target tracking application, what is the extent of crowding permissible in the selection region that can permit access and reporting of the desired target? This problem holds for the case of both crowding of multiple desired targets or a single desired target in a crowd of other objects. In order to detect more than one target, a widely used approach is to incorporate multiple transducers of the same type on board to indicate the presence of multiple targets [2], [11-12]. While the inclusion of multiple target-detectors on-board is a way to increase the detection capacity of the WSN, there is a clear difference between this method and the method of detection using an increased density of 'awake' nodes with a single target-detector on board. This can be illustrated by an example: in a habitat monitoring application to spot a certain species, two organisms of the same species in close proximity might register as one organism with the sensors in the nearby area. Unless the sensors are equipped with collocated multiple target detectors, an increased density of sensors might be less effective than a single sensor with efficient detecting/tracking abilities.

This brings us back to the problem of crowding. Since the probability of target detection across the deployment region is non-uniform and since activating multiple detectors on-board is not energy-efficient, one way to accommodate a sensitivity to acuity would be to develop intelligent networks that study the pattern of variation of the target and then perform adaptive density control.

3.2. Coding singularity

The coding singularity is relevant in understanding the coverage paradigm in WSNs, where coverage and connectivity are the primary factors in obtaining reliable network operation. In human cognition, the coding singularity limit refers to the constraint of selecting a given object or attributes of an object from the selection region. It differs from the acuity limit in that while the acuity limit focuses on the minimum spacing between items that allows access to individual items, coding singularity refers to the sparse description of the target in the selection region. Coding singularity in WSNs refers to the sensing resolution of the network which is defined by the resolution of the fundamental sensing unit: the sensor nodes. The area covered by a node that lies within its sensing radius is the finest level of detail that can be accessed by the sink for data processing.

Given this, the next question is: what should be the sensing resolution? The answer to this is application and objective dependent, although having data available at the finest resolution increases the reliability. This comes at a cost, since a high level of reliability requires a greater density of nodes sensing and transmitting data to the base station for further processing. Within this reliability constraint, the coding singularity poses two more issues. Firstly, there is the issue of what to transmit in a continuous sensing application like environment monitoring. Secondly, in an event-driven application like a target tracking/intrusion detection application, how do we recognize the target? Does merely increasing the density of 'awake' sensors guarantee an accurate response?

In a continuous sensing application, the uninterrupted nature of sensing and data dissemination has led to research into determining the subset of actual data that may be transmitted to the base station. Redundancy in deployment patterns has been exploited to reduce the transmission of redundant data due to spatial correlation in sensor locations or temporal correlation due to the pattern of variation in the sensed environment. While the coding singularity limit for attention in the neural system refers to the inability to process the features of more than one object in a selection region, this limit does not apply to WSNs. This is due to the presence of multiple transducers on board a sensor node that can sense multiple parameters of the sensed environment. However, coding singularity plays a role in information selection when the data processed at the base station is required to yield more information than merely the variation of the sensed parameters. Equivalently, this is a case of more unknowns than parameters, where the sensed parameters are processed to provide more information about the sensed field than can be obtained from transducer data in individual nodes. Coding singularity in a continuous sensing application is thus more relevant at the base station than at the nodes where the base station has to intelligently decide the amount of processing to be done on the gathered data to obtain relevant information. Alternately, in case of networks where nodes perform processing, it increases the complexity of determining what is relevant, since a node by itself has access only to the data within its sensing radius and to know the data from other nodes, it has to resort to increased inter-node communication which results in faster battery energy depletion and consequently affects network lifetime.

In response to the second issue, we recall that in the introduction, we mentioned the correlation between attention and performance. While increased attention improves performance, it does not hold true in the absence of a selection region. This is best illustrated in the case of a WSN deployed for target tracking or intruder detection application. If the features of the target are not provided to the network, there can be no awareness of the target even though all sensors are 'awake' and are transmitting gathered data to the sink. This holds also in case of inadequacy of the supplied features. If the target features are accurately outlined, it increases the efficiency of the target detection application in terms of decreasing/ eliminating the data propagation time from nodes to sink and the processing time at the sink to identify the target. Alternately, a faulty selection region that focuses on detecting objects other than the desired target have the same effect of resulting in loss of network resources such as battery power due to increased density of 'awake' sensors. The adequacy of supplied features acts like cues to the network to aid in efficiency of detection. The same argument can be used for continuous sensing scenario such as environment monitoring applications such in weather detection and temperature monitoring; however, the nodes do not have to perform the same level of processing as in detection applications to sense and report temperature. In other words, coding singularity is more relevant to tracking/detection applications with an emphasis on accurate selection. Figure 4 shows a cognitive WSN with the capacity, acuity and coding singularity limits that impact WSN performance.

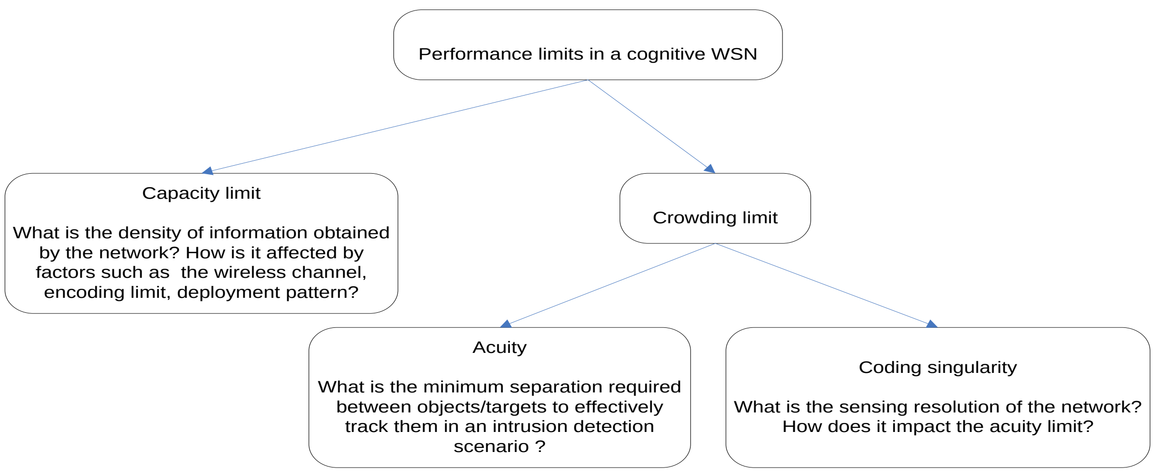


Figure 4. Guidelines for developing a cognitive WSN considering the capacity, acuity and coding singularity limits that impact WSN performance

3.3. Relationship between coding singularity and acuity

In this section, we illustrate the relationship between acuity and coding singularity limits for tracking application in WSNs. Acuity and coding singularity both derive from the issues of sensing resolution in the network. In the absence of a limit for coding singularity, the base station would have access to infinite amount of data obtained from the base station and not perform any compression to process the information from the raw data. The finest resolution of sensing would thus determine the quality of the sensing operation. However, in WSNs, clustering and in-network data processing performed at nodes allow for a certain leniency in estimating the information content from a given region of the deployment region. For example, spatial and temporal correlation from sensor locations and knowledge of variation in sensing field can be used to extrapolate the data from sensors that have been turned off due to power-saving mechanisms implemented at the nodes. The acuity limit is related to the crowding of targets in the selection region. While the region of interest (ROI) can be densely covered with an increase in the number of 'awake' sensors, they do not capture the amount of detail as fewer sensors with multiple transducers on-board. Thus similar to attention in neuroscience, the acuity limit for detection exists only because of the coding singularity that defines the sensing resolution of the network (due to coverage). However, the coding singularity does not determine the minimum separation between targets, i.e. limit for the acuity of detection.

4. Concluding remarks

The nature of WSN operation by distributed data gathering and processing in large-scale networks of nodes suggests that the primary goal of a WSN is to pay attention to the environment to sense the phenomenon of interest. The concept of attention in cognition can be leveraged to understand the nature of data gathering in WSNs. The knowledge of the limited nature of attention has led neuroscience research to explore among many avenues, the cognitive impact of limited attention. In this paper, we showed that the limits of capacity, acuity and coding singularity that limit attention in human cognition are also found in WSNs. In WSNs, these limits are manifested in the form of capacity of the network, ability for multiple target detection and sensing resolution of the network. We believe this framework of attention limits, which has been illustrated in this paper with the help of comparisons to the problems encountered in WSNs, will provide a unifying framework for studying the performance of WSNs. This study is worth pursuing in order to develop application-specific WSNs that do not just pay attention to the environment, but also adaptively learn to harness different 'types' of attention to provide the highest reliability of operation. The insights gained from an attention-oriented study can be used to develop self-organizing WSNs that allow for a combination of dynamic network topology, power management and routing techniques according to the variation of the sensing field. Our future work would involve analytical models of attention for reliable operation in WSNs.

5. References

[1] D. Marco, E. Duarte-Melo, M. Liu, D.L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data", Proceedings of the 2nd International Workshop on Information Processing in Sensor Networks, April 2003, pp: 1-16. Palo Alto, California.

[2] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, J. Anderson, "Wireless sensor networks for habitat monitoring", Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications, Sept. 2002, pp: 88-97, Atlanta, Georgia.

[3] M. Sims, J. Kurose, V. Lesser, "Streaming versus batch processing of sensor data in a hazardous weather detection system", Proceedings of the 2nd Annual IEEE Communications Society Conference on Sensors and AdHoc Communications and Networks, Sept. 2005, pp: 185-196. Santa Clara, California.

[4] M. Li, Y. Liu, "Underground structure monitoring with wireless sensor networks", Proceedings of the 6th International Workshop on Information Processing in Sensor Networks, Apr. 2007, pp: 69-78. Cambridge, Massachusetts.

[5] W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks", Proceedings of the 33rd International Conference on System Sciences, Jan. 2000, pp: 10-19. Maui, Hawaii.

[6] P. Cavanagh, "Attention routines and the architecture of selection", in M. I. Posner (editor). Cognitive neuroscience of attention, New York: Guilford Press, 2004. pp: 13-28.

[7] S. Haykin, "Cognitive radio: brain-empowered wireless communications", IEEE Journal on Selected Areas in Communications, Vol. 23, no. 2, Feb. 2005, pp: 201-220.

[8] T. Moscibroda, "The worst-case capacity of wireless sensor networks", Proceedings of the 6th International Workshop on Information Processing in Sensor Networks, Apr. 2007, pp: 1-10. Palo Alto, California.
[9] R. Machado, S. Tekinay, "Diffusion-based approach to deploying sensors to achieve coverage, connectivity and reliability", Proceedings of the 2nd International Conference for Advances in Wireless Sensor Networks, Aug. 2007, pp: 1-8. Philadelphia, Pennsylvania.
[10] L. Miao, H. Qi, F. Wang, "Biologically- inspired self-deployable heterogeneous mobile sensor networks", Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Aug. 2005, pp: 2363-2368. Alberta, Canada.
[11] P. Chen, O. Songhwai, M. Manzo, B. Sinopoli, C. Sharp, K. Whitehouse, et al, "Instrumenting wireless sensor networks for real-time surveillance", Proceedings of the IEEE International Conference on Robotics and Automation, May 2006, pp: 3128-3133. Orlando, Florida.
[12] G. Werner-Allen, J. Johnson, M. Ruiz, J. Lees, M. Welsh, "Monitoring volcanic eruptions with a wireless sensor network", Proceedings of the 2nd European Workshop on Wireless Sensor Networks, Jan. 2005, pp: 108-120. Istanbul, Turkey.

Authors

Renita Machado received her B.E. degree in Electronics and Communications Engineering with distinction from Manipal Institute of Technology, Manipal, India in 2004. She received her M.S. degree in electrical engineering from New Jersey Institute of Technology, NJ in 2005. She is currently pursuing her Ph.D. in Electrical Engineering while conducting research at the New Jersey Center for Wireless Networking and Internet Security (NJWINS lab) at New Jersey Institute of Technology. Her research interests are in mobile ad hoc networks, wireless sensor networks, pervasive computing and cognitive modeling for wireless networks.

Dr Tekinay has recently joined Ozyegin University as Vice Rector for Research and Technology. She continues to serve the US National Science Foundation where, in the last three years, she created two new programs while expanding the communications research program. She is the recipient of the NSF Director's Award for 'Excellence in Program Management' and her multidisciplinary program, 'Cyber-Enabled Discovery and Innovation', and the entire team received the Award for 'Collaborative Integrative Project'. She has been on faculty of the Electrical and Computer Engineering Department, New Jersey Institute of Technology since 1997 where she is currently an Associate Professor. She received NJIT's 'Excellence in Graduate Teaching Award' in 2003. Before joining academia, she worked at Bell Labs -Lucent Technologies and NORTEL. She holds eight patents. Dr Tekinay received her PhD in Electrical and Computer Engineering from George Mason University in 1994. She holds MS (1991) and BS (1989) degrees in Electrical Engineering from Bogazici University, Istanbul, Turkey.