

Learning In Wireless Sensor Networks for Environmental Monitoring

Dr M. Dash¹, Mrs. M. Balabantaray², Mr. D. P. Moharana³, Mrs. R. Mohanty⁴
(¹ Etc, Oec India)²etc,Oec India)³etc,Oec India)⁴etc,Oec India)

Abstract: This paper brings two main contributions to the use of learning techniques in a sensor network. First, an approach which combines time series prediction and model selection for reducing the amount of communication called adaptive model selection, is to let the sensors determine in an automated manner a prediction model that does not only fits their measurements, but that also reduces the amount of transmitted data. The second main contribution is the design of a distributed approach for modeling sensed data, based on the principal component analysis. We first show that the sensor measurements can be transformed along a routing tree in such a way that (i) most of the variability in the measurements is retained, and (ii) the network load sustained by sensor nodes is reduced and more evenly distributed.

Keywords: Learning techniques, sensor network, time series prediction, adaptive model selection.

I. Introduction

Wireless sensor networks form an emerging class of computing devices capable of observing the world with an unprecedented resolution, and promise to provide a revolutionary instrument for environmental monitoring. Such a network is composed of a collection of battery-operated wireless sensors, or sensor nodes, each of which is equipped with sensing, processing and wireless communication capabilities. Recent advances are in microelectronics and wireless technologies, wireless sensors are small in size, and can be deployed at low cost over different kinds of environments in order to monitor both over space and time the variations of physical quantities such as temperature, humidity, light, or sound. In environmental monitoring studies, many applications are expected to run unattended for months or years. Sensor nodes are however constrained by limited resources, particularly in terms of energy. Since communication is one order of magnitude more energy-consuming than processing, the design of data collection schemes that limit the amount of transmitted data is therefore recognized as a central issue for wireless sensor networks. An efficient way to address this challenge is to approximate, by means of mathematical models, the evolution of the measurements taken by sensors over space and/or time. Indeed, whenever a mathematical model may be used in place of the true measurements, significant gains in communications may be obtained by only transmitting the parameters of the model instead of the set of real measurements. Since in most cases there is little or no a priori information about the variations taken by sensor measurements, the models must be identified in an automated manner. This calls for the use of machine learning techniques, which allow to model the variations of future measurements on the basis of past measurements. In wireless sensor networks, communication is among the most energy-consuming task for a wireless sensor. This paper focuses on the design of learning techniques that trade data accuracy with communication by means of prediction models. Since sensor network measurements are very often correlated, the prediction models can often significantly reduce the communication while causing little loss in the accuracy of the measurements. Wireless sensor networks (WSN) form an emerging class of networks able to monitor environments with high spatiotemporal accuracy. The network is composed of tiny devices known as wireless sensors or motes, endowed with a microprocessor, a memory, a radio, a battery, and one or more sensors such as temperature, humidity, light or sound sensors [3]. The transmission of data from a WSN to an observer raises numerous issues: wireless sensors are constrained by limited resources, in terms of energy, network data throughput, and computational power. The communication module is a particularly constrained resource since the amount of data that can be routed out of the network is inherently limited by the network capacity. Also, wireless communication is an energy consuming task, identified in many situations as the primary factor of lifetime reduction [3]. The sensor network data are very often correlated both over space and time. Machine learning algorithms can be used to detect these redundancies, and to represent them by means of mathematical models. The use of mathematical models instead of the raw data can allow to substantially reduce the amount of data transmitted in the network, and thus to extend application lifetime. A typical scenario for a sensor network consists in placing a set of wireless sensors in an environment, such as a field, a forest or a town, and to use the collected measurements to monitor, detect, or track the evolution of a phenomenon over space and time. The network is usually connected to a base station by means of a routing tree, such as illustrated in Figure 1. The base station allows to centralize the data collected from the network, and acts as a gateway between the sensor

network and observers. It may be connected to the Internet, which allows the remote observation of the phenomenon monitored by the sensor network.

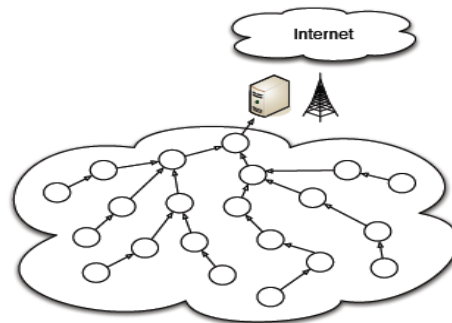


Figure 1. Multi-hop network architecture for an environmental monitoring application. A routing tree connects the sensor nodes to a base station, which connects the network to the Internet.

II. Wireless Sensor Networks

A Wireless Sensor Network is a collection of densely deployed autonomous devices, called sensor nodes, that gather environmental data with the help of sensors. The untethered nodes use radio communication to transmit sensor measurements to a terminal node, called the sink. The sink is the access point of the observer, who is able to process the distributed measurements and obtain useful information about the monitored environment. Sensor nodes communicate over a wireless medium, by using a multi-hop communication protocol that allows data packets to be forwarded by neighboring nodes to the sink. This concept is illustrated in Figure 2. The environmental or habitat monitoring is usually done over a long period of time, taking into account the latency requirements of the observer. The WSN can vary in size and topology, according to the purpose it serves. The sensor network is assumed to be homogeneous where nodes share a common communication medium (e.g. air, water, etc.). Further assume that the communication range is equal in size and strength for all nodes. They have a single Omni-directional antenna that can only broadcast a message, delivering it to all nodes in range. Sensor nodes can neither vary their transmission power, nor are they able to estimate their distance from the transmitting node by measuring the signal strength such features are not generally available in sensor nodes and therefore are not considered here. The motivation to use such simple devices is to reduce the overall cost of nodes and to keep our solution applicable to the most general sensor network.

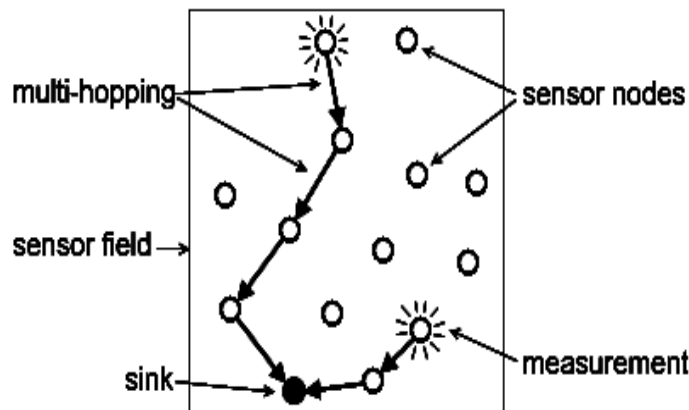


Figure 2. A Wireless Sensor Network using for a multi-hop communication protocol

III. Sensor Technology

A wireless sensor, or sensor node, is a device typically composed of a microprocessor, a memory, a radio transceiver, a power source and one or more sensors [3]. In many WSN applications, it is important that the measurements are geo localized. When sensors are randomly deployed, sensor nodes may also feature a geo positioning system (GPS) to obtain the location information. The schematic of a basic wireless sensor network devices is represented in Figure 3. The current generation of commercially available wireless sensor hardware have the size of a small wallet, and are designed mainly for experimental research purposes. The next generation of wireless sensors is well illustrated by the seminal smart dust project, which took place at the University of Berkeley between 1998 and 2001, and which led to the design of a laboratory prototype wireless sensor whose volume was about 5mm³. The design of viable millimeter scale wireless sensors is however still the subject of

research efforts. The resources available on a sensor node inevitably depend on its size. The need for experimenting WSN protocols, algorithms and applications in real-world conditions has led to the design of wireless sensor platforms for research. The University of Berkeley initiated such designs in 1999 with the WesC, a platform that approximated the functionalities envisioned by the Smart Dust project. This design was followed by the MICA, MICAz, MICA2DOT, and TelosB platforms, which have been the most widely used sensor nodes in academia for prototypical WSN deployments. The characteristics of these sensor nodes, also known as motes, are summarized in Figure3. These nodes use components-off-the-shelf (COTS) hardware instead of integrated silicon designs in order to allow easy customization of the boards, and to reduce the production costs.

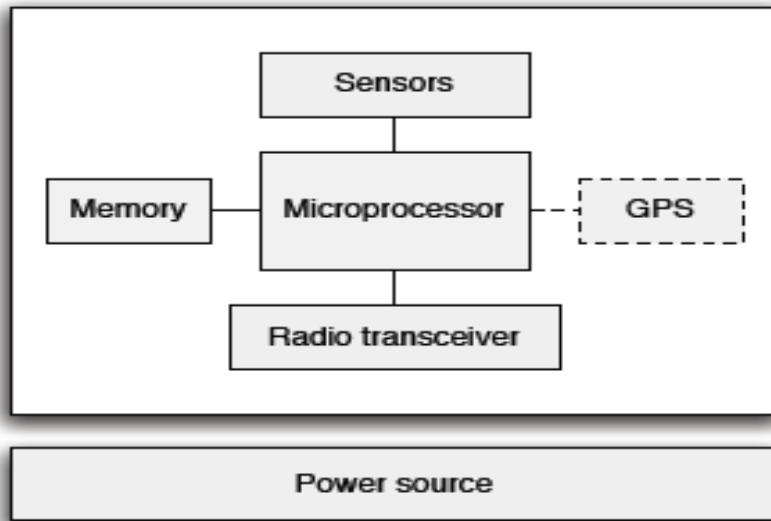


Figure 3 Wireless sensor network devices when sensors are randomly deployed, sensor nodes may also feature a geo positioning system (GPS) to obtain the location information.

Following the Smart Dust vision, shown in figure.4 these platforms run 8-bit microcontrollers and have a few tens of kilobytes of memory. Low data rate radios, ranging from a few tens to a few hundreds of kilobits per second, enable the communication. Sensors are either integrated (such as on the dot 2000 or the TelosB), or attached by means of a daughter board. Their size essentially depends on the batteries, typically a pair of AA cells. The integration of these COTS platforms in silicon would reduce their size to a few millimeter cube, as illustrated by the Spec platform, the silicon integrated counterpart of the MICA platform in the Smart Dust project. Such small scale wireless sensors are however still not viable, as a number of issues in hardware robustness and communication protocols must be further investigated.

Mote Type	WeC	René	René2	Dot	Mica	Mica2Dot	Mica 2	Telos
Year	1998	1999	2000	2000	2001	2002	2002	2004
Microcontroller								
Type	AT90LS8535		ATmega163		ATmega128		TI MSP430	
Program memory (KB)	8		16		128		48	
RAM (KB)	0.5		1		4		10	
Active Power (mW)	15		15		8		33	
Sleep Power (μ W)	45		45		75		75	
Wakeup Time (μ s)	1000		36		180		180	
Nonvolatile storage								
Chip	24LC256			AT45DB041B			ST M25P80	
Connection type	I ² C			SPI			SPI	
Size (KB)	32			512			1024	
Communication								
Radio	TR1000			TR1000		CC1000		CC2420
Data rate (kbps)	10			40		38.4		250
Modulation type	OOK			ASK		FSK		O-QPSK
Receive Power (mW)	9			12		29		38
Transmit Power at 0dBm (mW)	36			36		42		35
Power Consumption								
Minimum Operation (V)	2.7		2.7		2.7		1.8	
Total Active Power (mW)	24		27		44		89	
Programming and Sensor Interface								
Expansion	none	51-pin	51-pin	none	51-pin	19-pin	51-pin	16-pin
Communication	IEEE 1284 (programming) and RS232			(requires additional hardware)		USB		USB
Integrated Sensors	no	no	no	yes	no	no	no	yes

Figure 4 The family of Berkeley motes and their capabilities

IV. Learning Techniques In Sensor Networks

The primary interest of learning techniques in the field of sensor networks is to reduce the amount of energy consumed by sensor nodes by reducing the amount of communication in the network. The radio is indeed the most energy consuming component on a sensor node, and sensor nodes are battery operated. Reducing the amount of communication is therefore one of the main factors for extending the lifetime of the network, which is an important requirement of environmental monitoring applications where sensor deployments are expected to run for months or years. Learning techniques in sensor networks are attractive for the two following reasons. First, there usually exists a degree of redundancy in sensor network data. The redundancy stems from the fact that geographically close sensors are likely to collect similar measurements, and that measurements taken at two consecutive time instants are also likely to be similar. Learning allows to detect these redundancies, and to find mathematical models that represent the variations in sensor data in a more compact way. Second, it is rarely necessary to collect the exact measurements, and approximations can usually be tolerated by the observer. The lower the accuracy, the more compact the learning model, and therefore the higher the possible gains in communication. This makes possible the design of data collection strategies which trade accuracy for energy. Besides on its hardware, the energy consumption of a node is also dependent on its position in the WSN. Nodes, closer to the sink have to forward more messages and therefore need to listen more, while those far away from the sink could spend more time sleeping. For this reason, the behavior of agents cannot be the same for all (e.g. all listen and sleep the same amount of time in a frame). Each node needs to learn what behavior is energy efficient in the network. To achieve that, the nodes adopt an algorithm for optimization in order to improve the performance of the whole system. Each agent in the WSN uses a reinforcement learning (RL) algorithm to learn an optimal schedule (i.e. sleep duration in a frame) that will maximize the energy efficiency and minimize the latency of the system in a distributed manner. The main challenge in such a decentralized approach is to define a suitable reward function for the individual agents that will lead to an effective emergent behaviour as a group. Another challenge is that agents in a WSN can obtain only local information from surrounding nodes, due to their small transmission range.

V. Challenges In Applying Learning

A set of challenges are addressed and can be summarized as follow :

How to determine which model to choose?

The field of learning has a long history in computer science. As a result a wide range of techniques, such as linear models, neural network, regression trees or support vector machines, have been developed to tackle the problem of modeling the relationships existing in a set of data. The choice of a model that fits the data is however often difficult in practice, particularly when there is no a priori information on the relationships existing in the data. The limited processing resources of sensor nodes can furthermore exclude the use of computationally intensive learning techniques.

How to compute the parameters of a model in a distributed way?

When a model incorporates the spatial domain, this inevitably implies communication between sensor nodes in order to assess the relationships between the sensor measurements over space. The design of efficient distributed strategies to assess the spatial relationships is very challenging, and considerably reduces the range of techniques that can effectively be considered.

How to assess the ability of a model to properly extract the information of interest?

In practice, the accuracy of model can be assessed using different criteria, which depend on the application requirements. It can be for example the maximum deviation between the true measurements and the approximated measurements provided by the model, or the percentage of correct classification. An important issue in this respect consists in providing the observer with good estimates of the ability of the model to extract the requested information, and to ensure that this ability is maintained over time.

VI. Supervised Learning

In supervised learning. It introduces the general problem of learning with computers, and provides the general statistical framework that formalizes the actors of supervised learning tasks.

Learning “by heart” is probably one of the easiest thing for a computer. However, in most cases, the goal pursued is not to make the computer store all the examples provided, but rather to make it “understand” the underlying relationship that exists between the training examples provided. In particular, the goal of a learning procedure is most often to make good predictions, i.e., to properly guess the relationships existing between data other than those used as examples. This ability, called generalization, raises particularly tough challenges. An overview of the supervised learning setting is given in Figure 5.

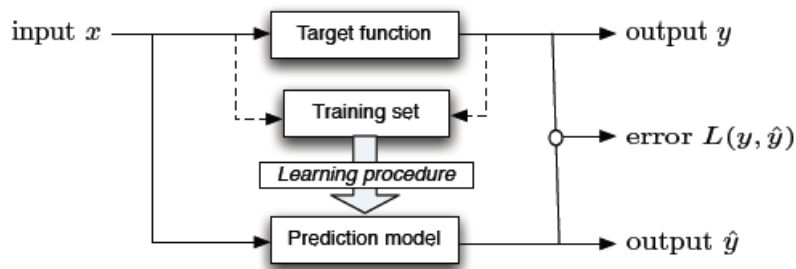


Figure 5 The supervised learning setting.

A training set made of input and output observations of the target operator is used by a learning procedure to identify a prediction model of the target operator. The goal is to minimize the error between the true output y and the prediction model \hat{y} , quantified by a loss function $L(y; \hat{y})$.

The data collection stage is an important step of a learning task, and the quality of the data obtained at this stage strongly impacts on the ability of the learning process to properly identify the hidden relationships. Indeed, no matter the efficiency of the search procedure in the space of models, the learning process is bound to fail if the relationships are overly complex between input and output variables. Different kind of data preprocessing techniques may be further applied to improve the efficiency of the learning phase. First, techniques such as data filtering, detection and removal of erroneous data (also referred to as outliers), and inference of missing data may be used to improve the quality of the data. Second, feature extraction techniques may be used to transform the original set of variables in a set of variables more relevant for the learning task.

The art on the use of learning techniques for reducing the amount of communication in sensor networks. Classifying these approaches in three main types, namely model driven, replicated models, and aggregative approaches, outlining for each of them their strengths and their limits. Table 1 gives a summary of the different learning schemes in terms of error type and highest network load.

Learning scheme	Error type	Highest network load
Model-driven acquisition	Probabilistic bounded or unbounded	$L_{\max}^{MD} \sim O(S_q)$
Replicated models	ϵ -bounded	$L_{\max}^{RM} \sim O(S)$
Aggregative approaches	Unbounded or ϵ -bounded	$L_{\max}^{DR} \sim O(p^2)$ $L_{\max}^{DR_{\text{check}}} \sim O(p^2 + S)$

Table 1: Comparison of the performances of the different modeling approaches

Approaches based on model-driven acquisition reduce the highest network load to S_q , i.e., the number of sensors whose measurements are effectively collected. The main characteristic of these approaches is that part of the network can remain in the idle mode. Model-driven data-acquisition therefore not only reduces the highest network load, but also allows to reduce to a negligible level the energy consumption of the sensor nodes not queried. The subset of sensor nodes whose measurements are collected could be changed over time to distribute the energy consumption. Indeed, there exists in most cases different pairs of set of queried and predicted sensors for which the observer's accuracy requirements can be satisfied. The main characteristic of replicated models is to guarantee bounded prediction errors, even in the case of unexpected events. These approaches however require the sensor nodes to take their measurements at every epoch, so that they can be compared with the predicted measurements. The energy savings depend on the frequency of updates of the model. Finally, aggregative approaches lead to either unbounded or bounded error. The type of error depends on whether aggregates obtained at the base station are communicated to sensor nodes. If no check is made against the true measurement, the highest network load is reduced to the number of aggregates collected.

VII. Machine Learning For Energy-Efficient Monitoring

Among the different tasks that must be performed by a sensor node, the radio use is by far the most expensive in terms of energy consumption. On a typical sensor node such as the Telos mote, the radio expends one order of magnitude more energy than the CPU over an equivalent length of time [15]. The lifetime of a sensor node therefore highly depends on how frequently the radio is used. A central issue in sensor networks thus consists in reducing the amount of communication in the network, while providing the observer with the

requested information. Machine learning techniques in computer science consist in inferring a prediction model, on the basis of a set of observations. The model is in most cases a parametric function, which allows to predict the value of a variable, called output, given a set of other variables, called inputs. In the domain of sensor networks, the use of learning techniques is attractive for a number of reasons. First, sensor data are very often correlated both over space and time. For example, outdoor temperatures typically follow consistent diurnal and seasonal patterns, and at any moment in time, their variations are unlikely to vary greatly within a local region.

Learning techniques can be used to detect and model the relationships existing between sensor data, both at the temporal and spatial scales. The use of the models in place of the raw measurements can greatly reduce the amount of data transmitted in the network. Second, approximations of the sensor measurements can usually be tolerated by the observer. For example, in climatic studies on plant growth, it is often sufficient to collect temperature and humidity measurements within 0.5 and 2% of the true quantities [19]. One feature of learning techniques is that they can adapt to different levels of accuracy, by using models with varying degree of complexity. The complexity of a model typically depends on its number of parameters. When approximations can be tolerated, simple models with fewer parameters may be used to further reduce the amount of communication. Finally, the interest of the observer is not necessarily in the measurements, but in higher level information such as the types or number of animals present in the monitored environment for example. The extraction of such high-level information typically requires to fuse and transform the measurements from different sensors. Learning techniques provide an effective way to determine how the measurements can be combined in order to infer the requested information. In some cases, it is possible to combine the measurements within the network, thus avoiding the transmission of all the measurements to the base station. These strategies, referred to as in-network data aggregation, are among the most promising ones for efficiently extracting information from a sensor network [25]. The temperature variations can vary greatly across the office. However, they locally present linear trends, which can be effectively captured by linear models. Each model is here a function of spatial coordinates that are linearly mapped to temperature measurements. Each model requires only three parameters to represent the variations within a region, and provides the observer not only with approximations of the measurements at the sensor's locations, but also at any Location within a given region.

VIII. Learning With Invisible Media

Wireless sensor networks are the latest trend of technological advances which were well predicted by Gordon Moore, the co-founder of Intel, some 45 years ago. Known as the Moore's law in the field of computer science, the law states that the number of transistors that can be placed inexpensively on an integrated circuit doubles every two years. This law has been remarkably verified since it was stated, and computing devices have become more and more part of our daily lives. The Eniac was created in 1946, at the time referred to as the "giant brain", and had a volume of 63m² and consumed 150kW of power. Advances in electronics have allowed to reduce the price and volume of computational devices, which made their use possible to a wider audience. The 90s definitely made a corner in the history of computer science, by allowing many people to have their own Personal Computer, which could be used as a new communication device thanks to Internet. Cell phones have since then allowed to make communication possible anywhere and anytime, both using cellular networks and the Internet.

IX. Conclusion

Wireless sensor networks nowadays represent what will probably be tomorrow's part of our daily lives. WSNs allow to extend Internet to the physical world, providing real-time information about physical phenomena occurring in an environment.

This paper brought additional results to the feasibility of such real-time monitoring scenarios. In particular, the learning techniques could be efficiently used to reduce the communication among sensor nodes. The challenges brought by these new technologies are still numerous, and we believe that intelligent data processing techniques can address some of the most important. Data is at the heart of these sensing systems, and is what actually matters to the observer. The exponential amounts of data that such systems will generate in the coming years cannot however be made readily visible to the observer. Rather, data fusion and learning techniques must be further developed to allow collaborative and intelligent processing of the data within the network.

Reference

- [1] D.W. Aha. *Lazy learning*. Kluwer Academic Publishers Norwell, MA, USA, 1997.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. *Wireless Sensor Networks: A Survey*. *Computer Networks*, 38(4):393–422, 2002.
- [3] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. *Wireless sensor networks: a survey*. *Computer Networks*, 38(4):393–422, 2002.
- [4] ST Alexander. *Adaptive Signal Processing: Theory and Applications*. Springer-Verlag New York, Inc. New York, NY, USA, 1986.

- [5] Z. Bai et al. *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. Society for Industrial and Applied Mathematics, 2000.
- [6] M. Birattari, G. Bontempi, and H. Bersini. Lazy learning meets the recursive least squares algorithm. *Advances in Neural Information Processing Systems*, pages 375–381, 1999.
- [7] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, USA, 1995.
- [8] G. Bontempi. *Local learning techniques for modeling prediction and control*. Bruxelles:IRIDIA-Universite Libre de Bruxelles, 1999.
- [9] G. Bontempi and Y. Le Borgne. An adaptive modular approach to the mining of sensor network data. In *Proceedings of the Workshop on Data Mining in Sensor Networks, SIAM SDM*, pages 3–9, Philadelphia, PA, 2005. SIAM Press.
- [10] L. Breiman. *Classification and Regression Trees*. Chapman & Hall/CRC, 1998.
- [11] P.J. Brockwell and R.A. Davis. *Introduction to time series and forecasting*. Springer, 2002.
- [12] P. Buonadonna, D. Gay, J. M. Hellerstein, W. Hong, and S. Madden. TASK: Sensor Network in a Box. In *Proceedings of the 2nd IEEE European Workshop on Wireless Sensor Networks and Applications (EWSN'05)*, Istanbul, Turkey, February 2005.
- [13] N. Burri and R. Wattenhofer. Dozer: ultra-low power data gathering in sensor networks. In *Proceedings of the 6th international conference on Information processing in sensor networks*, pages 450–459. ACM Press, 2007.
- [14] C. Chatfield. *Time-series forecasting*. Chapman & Hall/CRC, 2001. [15] D. Chu, A. Deshpande, J.M. Hellerstein, and W. Hong. Approximate data collection in sensor networks using probabilistic models. In *International Conference on Data Engineering (ICDE)*, 2006.
- [16] A. Cichocki and S. Amari. *Adaptive blind signal and image processing*. Wiley New York, 2002.
- [17] N. Cristianini and J. Shawe-Taylor. *An introduction to support Vector Machines: and other kernel-based learning methods*. Cambridge University Press New York, NY, USA, 1999.
- [18] Intel Research Laboratory Data. Project Website. berkeley.intel-research.net/labdata/, 2003.
- [19] A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein, and W. Hong. Model-Driven Data Acquisition in Sensor Networks. In *Proceedings of the 30th Very Large Data Base Conference (VLDB'04)*, Toronto, Canada, 2004.
- [20] A. Deshpande, C. Guestrin, S.R. Madden, J.M. Hellerstein, and W. Hong. Modelbased approximate querying in sensor networks. *The VLDB Journal The International Journal on Very Large Data Bases*, 14(4):417–443, 2005.
- [21] J.M. Dricot, M. Van Der Haegen, Y. Le Borgne, and G. Bontempi. A modular framework for user localization and tracking using machine learning techniques in wireless sensor networks. In *Proceedings of the 8th IEEE Conference on Sensors*, pages 1088–1091, 2008.
- [22] J.M. Dricot, M. Van Der Haegen, Y. Le Borgne, and G. Bontempi. Performance evaluation of machine learning technique for the localization of users in wireless sensor networks. In L. Wehenkel, P. Geurts, and R. Marée, editors, *Proceedings of the BENELEARN Machine Learning Conference*, pages 93–94, 2008.
- [23] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern classification*. Wiley New York, 2001.
- [24] D. Estrin, L. Girod, G. Pottie, and M. Srivastava. Instrumenting the world with wireless sensor networks. In *Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on*, volume 4, 2001.
- [25] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi. In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey. *Wireless communications*, 14(2):70, 2007.
- [26] D. Gay, P. Levis, R. Von Behren, M. Welsh, E. Brewer, and D. Culler. The nesC language: A holistic approach to networked embedded systems. *Sigplan Notices*, 38(5):1–11, 2003.
- [27] J. Gehrke and S. Madden. Query processing in sensor networks. *Pervasive Computing, IEEE*, 3(1):46–55, 2004.
- [28] G.H. Golub and C.F. Van Loan. *Matrix computations*. Johns Hopkins University Press, 1996.
- [29] Good Food EU Integrated Project: Food Safety and Quality Monitoring with Microsystems. *Sensor Network in a Vineyard*. www3.unifi.it/midra/goodfood/, 2007.