Energy-Efficient Sensor Data Gathering in Wireless Sensor Networks

Ruqiang Yan*, Zhaoyan Fan[1], Robert X. Gao[1] and Hanghang Sun

Southeast University, No. 2 Sipailou, Nanjing, 210096, China
[1]University of Connecticut, 191 Auditorium Road, Storrs, CT 06269, USA

In this paper, we present the design and realization of energy-efficient sensor data-gathering schemes used in wireless sensor networks (WSNs) for monitoring engineering systems. A key issue in the design of WSNs is to establish a reliable assessment scheme by combining data acquired from each individual sensor into a single wireless sensor network. However, as the size of the network rapidly grows, aggregating information made by all the sensors becomes computationally intractable, causing increased energy consumption. Hence, it becomes critical that the WSN be sectioned to allow for computational efficiency while reducing the overall data communication requirements for the purpose of energy saving. Different data-gathering schemes have been investigated in this study for energy-efficient sensing. In order to automatically perform effective data gathering in such a network, the hardware and software of the sensor node are designed to meet the requirement of optimizing energy efficiency. It was found that the overall energy cost of the system is considerably less for a sectioned WSN than the conventional centralized network structure.

1. Introduction

Recent advancements in system miniaturization, wireless communication, and on-chip signal processing have enabled the development of low-cost, low-power, intelligent multifunctional sensors.[1,2] Such types of sensor generally consist of data acquisition, data processing, and communication components, and can be employed in various engineering application areas, such as aircraft structural health monitoring,[3] machine condition-based maintenance,[4] industrial electric systems monitoring,[5]

---

and environmental sensing.[6]   However, effective energy/power utilization remains a practical challenge for the full deployment of a wireless sensor network (WSN).

Prior research has focused on various issues related to power management, such as self-powered sensing using energy extracted from the manufacturing processes being monitored and the ambience.[7–9] Also studied was how to improve the energy efficiency so that a fixed amount of power supply (*e.g.*, a set of battery) can last for a maximum length of time.   This includes adaptive sampling,[10,11] dynamic voltage scaling,[12,13] multihop routing,[14,15] low-power protocol operation,[16,17] and data gathering.[18,19] Among various data-gathering algorithms, one of the solutions is to section the sensor network into several subnetworks and communicate only high-level inferences between the subnetworks,[20] which are divided on the basis of the logic structure of multiple sectioned Bayesian networks.[21]

Inspired by these research efforts, in this paper, we present the design and realization of an energy-efficient data-gathering scheme for utilization in WSNs, where the sensors are assumed to be able to communicate with each other, possess decision-making capabilities, and alter their performance based on the observed system state. Furthermore, the hardware and software for applying the proposed data-gathering scheme in machine monitoring application are designed and evaluated for optimized energy efficiency.   The remainder of this paper is organized as follows.   Section 2 introduces the candidate data-gathering schemes and the associated energy consumptions in WSNs. The design of the hardware and software of the sensor is described in § 3, based on the consideration of energy efficiency.   Then, the implementation and testing of a WSN is described in § 4, and conclusions are drawn in § 5.

2.   Data Gathering Schemes of Wireless Sensor Network

2.1   *Energy consumption calculation*

Generally, data computation and communication are the two major energy consumers in the WSN, which provide local data processing and intersensor information exchange, respectively.   Given a fixed sampling rate and data length, the energy cost for data computation can be calculated as

$$E_{comp} = N_{cyc} \cdot T_c \cdot V_c \cdot I_c, \tag{1}$$

in which N_{cyc} is the number of cycles needed for signal processing, T_c is the machine cycle time, V_c is the working voltage, and I_c is the current.
For two sensors to communicate, the energy consumption needed for data transmission can be expressed as [12]

$$E_{Tx} = E_{e\_tx} \cdot k + \varepsilon_{amp} \cdot d^\alpha \cdot k, \tag{2}$$

where k is the number of transmitted data bits, α is a factor valued from 2 to 5, depending on the environment of wireless transmission, d is the distance between two sensory nodes, ε_{amp} (J/b/m^2) is the amplification coefficient to satisfy a minimum bit error rate to

ensure reliable reception at the receiver, and $E_{\text{e\_tx}}$ (J/b) is the energy dissipated to operate the transceiver, which is given as

$$E_{\text{e\_tx}} = V_{\text{cc}} \cdot I_{\text{TP}} / K_{\text{data\_rate}}, \tag{3}$$

where V_{cc} denotes the working voltage, I_{TP} denotes the current for transmission, and $K_{\text{data\_rate}}$ denotes the data transmission rate.

The energy consumed for receiving a data stream can be expressed as

$$E_{\text{Rx}} = E_{\text{e\_rx}} \cdot k. \tag{4}$$

Equation (2) shows that for a fixed distance, the energy consumed is proportional to the length of the data bits. On the other hand, the longer the distance between two sensory nodes, the more energy will be consumed.

2.2   *Data gathering schemes*

Several data gathering schemes that have implications on the overall energy consumption of a WSN have been conceived. For simplicity, all the sensors in the network are assumed to be identical. The various data gathering schemes[20] are illustrated as follows:

Scheme 1: Raw data points are transmitted to the central monitoring unit (CMU) from each individual sensor.

Scheme 2: Only physical features extracted from the raw data for each sensor are transmitted to the CMU.

Scheme 3: The sensors are grouped into different segments, and the raw data from each sensor are transmitted to the corresponding segment head. After extracting relevant features, a data fusion operation is implemented by the segment head, and then the fusion information is transmitted to the CMU.

Scheme 4: The individual sensors are sectioned into different segments. Each sensor performs localized data processing, and only the features extracted from the raw data from each sensor are transmitted to the corresponding segment head. After data fusion, the segment head transmits the fused information to the CMU.

As an example, Fig. 1 shows the four data gathering schemes being applied on the platform of a machine monitoring test bed in which time- and frequency-domain features such as root mean square (RMS) value and ball pass frequency on the inner/outer raceway (BPFI/BPFO) extracted from bearing vibration signals are used to indicate machine status and estimate potential failures caused by mechanical defects.[22] As shown in Fig. 1(a), when working as scheme 1, all the bearing sensors transmit raw data and all the computation works were done by the CMU. In scheme 2, as shown by Fig. 1(b), the computation for feature extraction was assigned to each bearing sensor, so that only features like RMS and BPFO were transmitted to the CMU instead of all the raw data. By using the structure of a sectioned sensor network, Figs. 1(c) and 1(d) show the sectioned network with a segment head used as data relay and local section controller. In scheme 3, the feature extraction was only assigned to the segment heads, so the raw

(a)   (b)

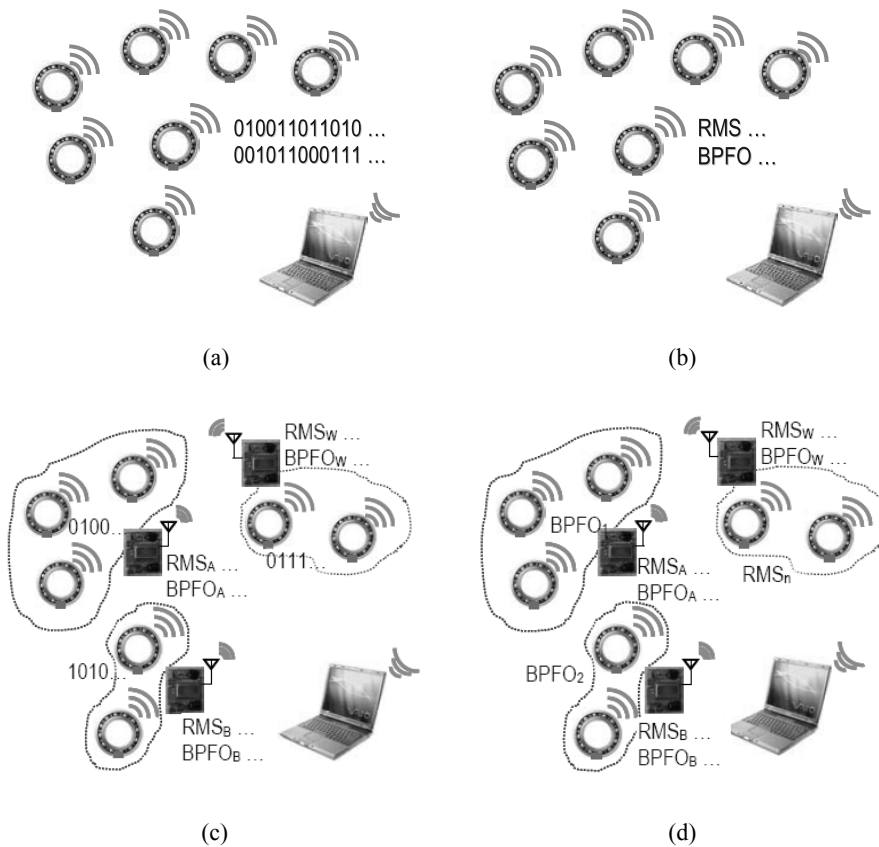(c)   (d)

Fig. 1.   Energy-efficient data gathering schemes used in sensor network. (a) Scheme 1, (b) Scheme 2, (c) Scheme 3, and (d) Scheme 4.

data still remained in the communication within the local sections. While in scheme 4, feature extraction was completed on each sensor. To be different from scheme 2, all the features are further combined by the segment heads for reduced communication load before forwarding the information to the CMU.

3.   Hardware and Software Design

To realize energy-efficient data gathering in the sectioned sensor network, the sensor system architecture is designed and illustrated in Fig. 2. Two major hardware functional modules, namely, computation and communication, are required to perform effective data collection, processing, and communication. Controlled by a power management software that is embedded with the computation module, the four data gathering schemes are performed to achieve minimum energy consumption of the WSN. A major focus of
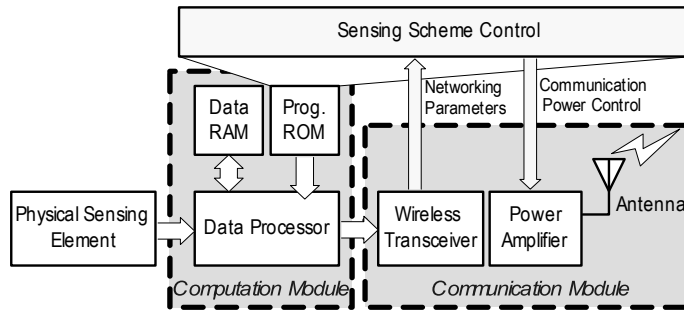
Fig. 2.   System architecture for sensors.

the sensor system design is to optimize both the hardware and software to be compatible with all the data gathering schemes, while keeping the system scale to a minimum for energy saving and prolonged battery life.

3.1   *Computation module*

Data computation, including local feature extraction and data fusion, is employed by individual sensors to reduce the length of transmitted data and save energy. For example, considering the scenario where vibration sensors are used for bearing health monitoring, the signal is often sampled at a rate above 5 kHz with a minimum window size of 1 s, as shown in Fig. 3(a), to cover the major characteristic frequency components like BPFO, BPFI, and their resonances with an acceptable resolution in frequency-domain analysis.[16] This translates to a length of 5 kB of raw data gathered in each sample period. Compared with various feature extraction techniques developed in the time and frequency domains, analytical and experimental studies[23] have demonstrated that the Discrete Harmonic Wavelet Packet Transform (DHWPT) is both effective and efficient in revealing the time-frequency characteristic of defect-induced vibration signals, in particular, in the high-frequency region (*e.g.*, over 10 kHz), as indicated by the columns in Fig. 3(b). Through a series of conversions using Fast Fourier Transform (FFT) and Inverse Fast Fourier Transform (IFFT), the raw data samples are divided into 16 subfrequency bands and the energy of each subband is calculated as the features. It has been proved experimentally[23] that the statistic distribution of these features effectively indicates the status of defect and successfully predicts the bearings' lifetime. Comparison between Figs. 3(a) and 3(b) shows a 1000-fold reduction in data length after local feature extraction using DHWPT. In this case, it is possible to achieve energy saving in data transmission, while reserving the defect-related information conveyed by the signals.

An important aspect pertaining to the computation module design is the effect of data length on the use of memory size for digital signal processing algorithms. Consider the DHWPT method used for feature extraction; the basic operation composed of FFT or
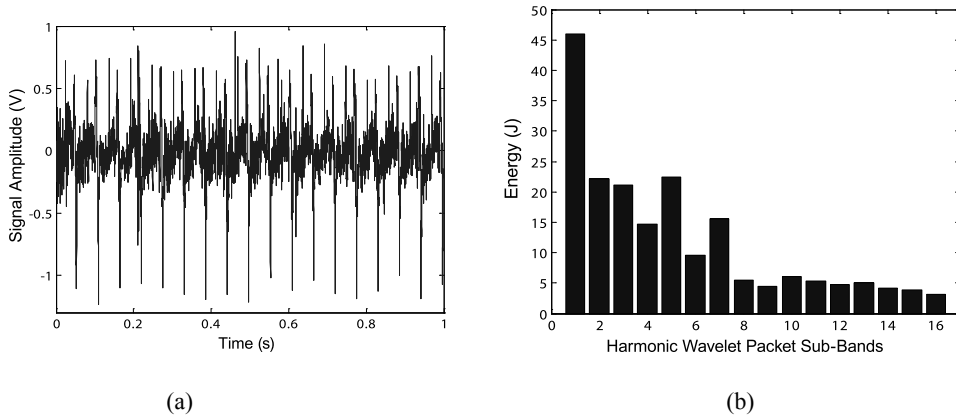
Fig. 3.   Time-frequency decomposition of a bearing vibration signal based on DHWPT. (a) Raw data from vibration sensors and (b) Features extracted by the sensory node.

IFFT is used to convert data from time domain to frequency domain and vice versa. Both FFT and IFFT algorithms require the same amount of RAM as the length of input data to buffer the temporary and final results. Given that the raw vibration data containing 8 bits is sampled at f_s Hz in a period of t_s seconds (corresponding to a resolution of $1/T_s$ Hz in spectrum analysis), the length of raw data is calculated as $N_s = f_s \cdot T_s$. Implementing the FFT algorithm requires choosing the input data length as a power of 2: $N_{FFT} = 2^p$, where $p = \lceil \log_2 (N_S) \rceil$ denotes the calculation of the minimum integer that is not smaller than $\log_2(N_S)$. A total number of $(N_{FFT} - N_S)$ zeros is appended to fill the increased data length by the zero padding technique. As an example, in cases where data is sampled at 5 kHz with 1 s period, the minimum required RAM size is calculated as 8192 bytes and 3192 zeros are added.

3.2   *Communication module*

In the four data gathering schemes designed for the WSN, it is assumed that the transmission power is minimized to ensure reliable reception at the receiver, according to the communication distance between sensors. Hence, the awareness of communication power, as well as adjustability of the transmitter output power, becomes critical to perform these schemes using hardware devices. By assuming a unit signal gain provided by antennas, the output power of the communication module is dominated by the consumption of the power amplifier. To transmit 1 bit to the receiver, the output power and associated received power are expressed as

$$P_{TX} = (\varepsilon_{amp} \cdot R) \cdot \hat{d}^{\alpha}, \tag{5}$$

$$P_{RX} = P_{TX}/d^\alpha = (\varepsilon_{amp} \cdot R) \cdot (\hat{d}^\alpha/d^\alpha) = P_S \cdot (\hat{d}^\alpha/d^\alpha), \qquad (6)$$

where R denotes the data transmission rate, $\hat{d}$ and d are the estimated and actual transmission distances between the transmitter and receiver, respectively, and $P_S = \varepsilon_{amp} \cdot R$ is the receiver sensitivity denoting the minimum signal power that the receiver can discern. From eq. (6), it is shown that if the estimated distance $\hat{d} < d$, then the received signal cannot be identified and the communication between sensors fails. On the other hand, if $\hat{d} > d$ (over the estimation), which shows a received power higher than sensitivity, a portion of the transmission energy will be lost on the propagation path while it does not affect the results of the signal reception. In this case, the energy efficiency problem was translated to the effective estimation of the communication distance between sensors.

　　Since all the sensors are equipped with both transmitting and receiving capabilities and the received signal power can be measured using hardware through receive signal strength indication (RSSI), as shown in Fig. 4, the distance estimation can be performed using a set of handshaking operations between the transmitter and receiver. By using microcontroller MC56F8323[24] as the computation and control unit, the distance can be estimated by a software-controlled procedure as shown in Fig. 5. For example, consider the case where scheme 4 is applied to local data transmission from a sensor to its segment head. In the first step, the segment head broadcasts a Test Code (such as a binary signal of 010101…, which has a duty cycle of 50%) with maximum transmit power P_{TXmax} to all the sensors in the local sectioned network. By measuring the received power $P_{RX(SHK)}$ on each sensor, the distance to the segment head can be calculated as

$$\hat{d}^\alpha \geq P_{RX(SHK)} / P_{TXmax} \quad . \qquad (7)$$

　　For a physical RSSI circuit model as shown in Fig. 4, the received signal power P_{RX} is translated to the voltage value, V_{RSSI}, through a high-pass filter, signal rectifier, and then
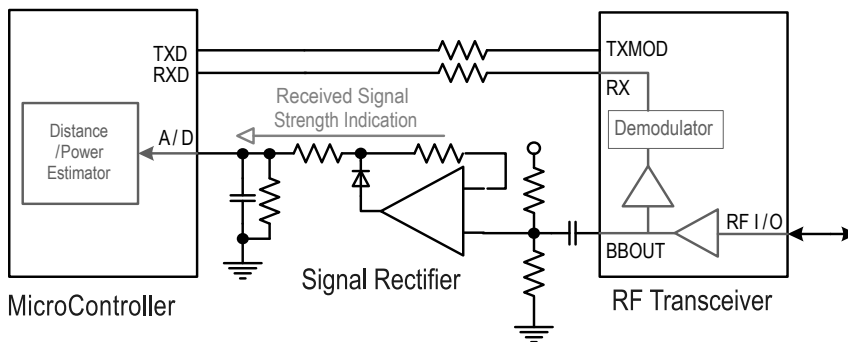

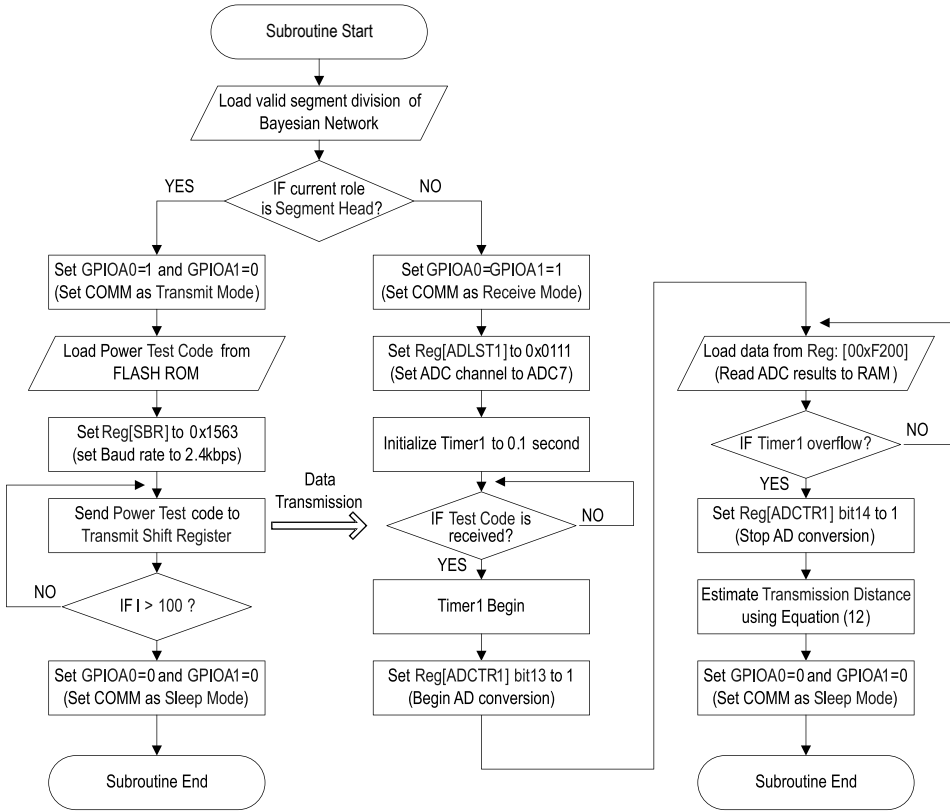Fig. 4.　Circuit design for receiver energy measurement and distance estimation.

Fig. 5.   Flow chart of communication distance estimation.

a low-pass filter. When the test code broadcasted by the transmitter (either segment head or CMU) has a fixed duty cycle, the output RSSI voltage using transceiver TR1001[25] is proportional to the received power at a constant rate of q (q = 10 mV/dB for duty cycle = 50%). By substituting this relationship to eq. (7), the estimated distance is expressed as a function of the RSSI voltage

$$\hat{d}^{\alpha} \geq 10^{\frac{V_{\mathrm{RSSI}}}{10\,q}} / P_{\mathrm{TXmax}}, \tag{8}$$

where V_{RSSI} is calculated with the unit of mV, q is calculated in mV/dB, and P_{TXmax} is calculated in mW. Hence, by minimizing the estimated distance for data transmission, the minimum required power to ensure data communication is expressed as

$$P_{\text{TX}} = P_{\text{S}} \cdot 10^{\frac{V_{\text{RSSI}}}{10 \cdot q}} / P_{\text{TXmax}}. \tag{9}$$

To always perform the minimized energy consumption on the communication module, an output power control circuit is designed, as illustrated in Fig. 6. To boost the transmission power for long-distance transmission, a power amplifier (PA) is added between the transceiver and antenna. The RF switch provides a secure swap between transmit and receive modes. The output power of the communication module is determined by the first-level amplifier of PA, which is controlled by the analog input (GC). Because the relationship between transmission power and distance is not linear, as indicated in eq. (2), it is impossible to control the GC to fit the distance with the equivalent resolution in the full scale by using a linear-adjusted D/A output. As a solution, a logarithmic amplifier is employed as a nonlinear translator from the D/A output of the microcontroller to the GC of the power amplifier. In this case, the communication energy in the WSN can be fully digitized by the hardware as well as be successfully estimated and controlled by the embedded software.

3.3 *Sensing scheme control*

Although communication and computation energy are individually minimized by the associated functional modules on each sensor, by using the appropriate data gathering scheme for the whole WSN, the total energy consumption can be further reduced. For example, in cases where sensors' computation energy dominates the system consumption, it is more efficient to choose scheme 1 (where there is no computation work assigned to
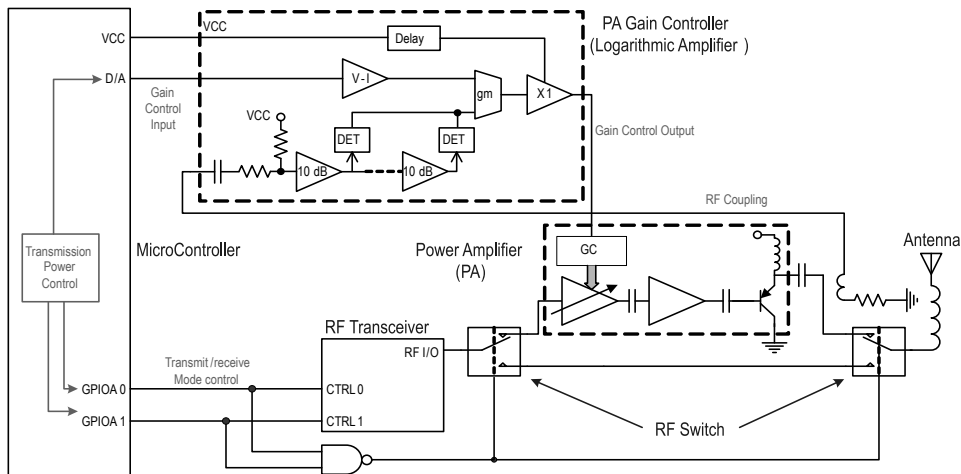

Fig. 6.   Circuit diagram of transmission power control mechanism for communication module.

sensors) to reduce the computation energy by increasing the communication load. On the other hand, if the communication energy consumed in the network is much higher than the computation energy, energy saving can be achieved by selecting schemes with a higher computation load assigned to sensors (scheme 2 and 4).

Figure 7 shows the flow charts for gathering energy-related information, calculating the potential energy consumption, and switching of the sensor's settings to perform the efficient sensing schemes. By assuming a homogeneous hardware structure for all the sensors, as shown in Fig. 7, the hardware-related parameters like ε_{amp}, E_{e_tx}, and E_{cmp} are predefined and stored into the ROMs on CMU, while variables representing distances between sensors/segment heads and CMU are aggregated by broadcasting a request to
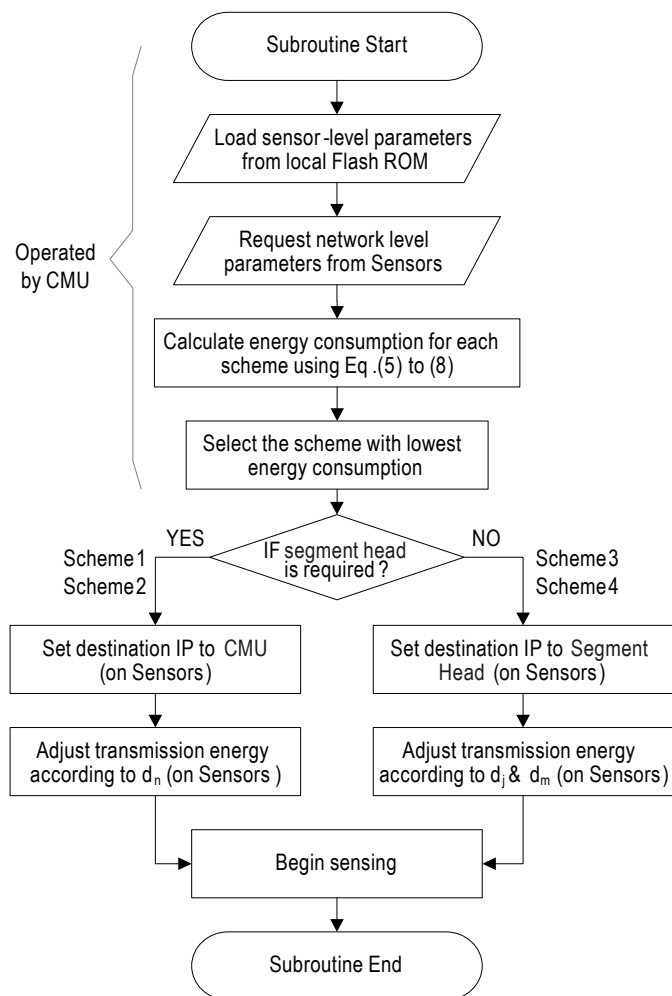

Fig. 7.   Flow chart for scheme selection steps on sensors and CMU.

all the sensors. After all the information is collected, CMU estimates the overall energy consumption for each of the schemes using eqs. (5) to (8), and makes a decision to choose the one with the best energy efficiency as the current data gathering scheme.

4.   Implementation and Test

Based on the presented design issues on energy-efficient hardware and software in section 3, a sensor platform for vibration-based machine health monitoring was implemented by using Kistler 8694M1, Motorola MC56F8323, and RFM TR1001 as for physical sensor, computation, and communication modules, respectively, as shown in Fig. 8. The control software developed for vibration signal processing and data fusing/inference combined with scheme selection was developed and evaluated on the Code Warrior environment.

An investigation was then conducted to evaluate the impact of various data gathering schemes on the energy efficiency of the WSN system (structure is shown in Fig. 9) on the basis of the parameters of the designed hardware and software platform (as listed in Table 1[24,25]). According to the machine health monitoring system, the DHWPT algorithm is applied to the local feature extraction. Assuming a sampling rate of 5 kHz, and an input data length of 8192 points from each sensor, the number of machine cycles for implementing the DHWPT algorithm is calculated as 7767712 cycles[23] and the number of machine cycles for data fusion/inference is counted to be twice that of the local feature extraction. Figure 10 illustrates the energy consumption of each scheme as the distances between sensors and CMU changes in the network. Basically, the energy consumption is calculated by using eqs. (1)–(4). The voltage and current parameters listed in Table 1 are used to estimate the computation energy cost, while the transmission energy cost is calculated using the working voltage and current being measured through a customized circuit.
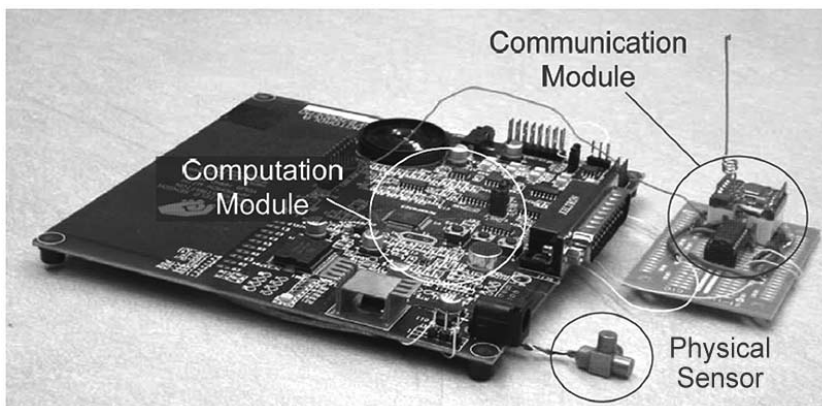

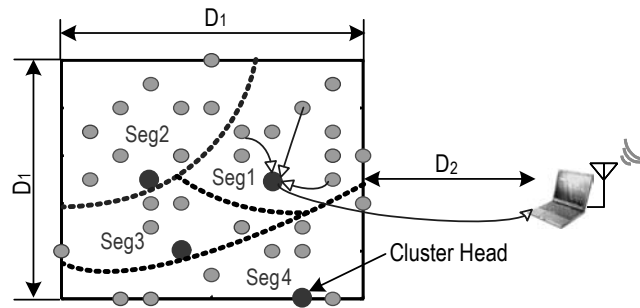Fig. 8.   Sensor model developed for energy-efficient machine health monitoring.
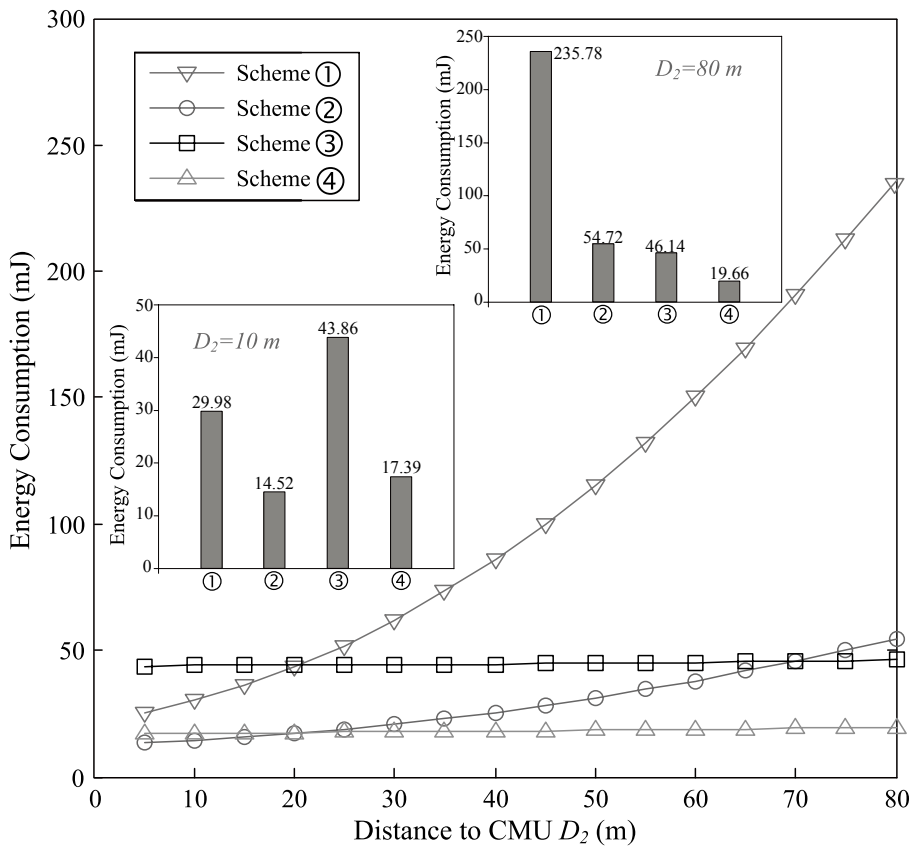
Fig. 9.   Sectioned sensor network structure.

Fig. 10. Energy consumption of the test network structure.

Table 1
Parameters for energy calculation.

Parameter	V_c (V)	I_c (mA)	T_c (s)	E_{e_tx} (nJ/b)	ε_{amp} (pJ/b/m^2)
Value	3	110	$1/(60 \times 10^6)$	6535	8934

It can be seen that, for a fixed distance D_1, the energy consumption increases as D_2 increases for each data gathering scheme. Furthermore, when D_2 is short (*e.g.*, D_2 = 10 m), the energy consumption of a sectioned network (corresponding to Schemes 3 and 4) is higher than that of the conventional network (corresponding to Schemes 1 and 2), for both the raw data transmission and feature transmission. This is due to the fact that the energy dissipated to run the transmitter is dominant for short-distance data communication. Furthermore, for the sectioned network, energy will also be consumed by the segment heads to receive the data from individual sensors, thus leading to higher energy consumption for Scheme 3 than for Scheme 1, as well as for Scheme 4 than for Scheme 2. However, with the increase in the distance from the WSN to the CMU, the energy needed to amplify the transmission to ensure communication between the transmitter and the receiver becomes dominant. Since the related power is proportional to the square of the distance of data transmission, a sectioned network is more energy efficient than a conventional network. It was found that both schemes used in the sectioned network were more energy efficient than those used in a nonsectioned network, when D_2 is large (*e.g.*, D_2 = 80 m).

5.   Conclusions

Effective diagnosis of the health status of a system using a large-scale WSN requires both accurate data gathering and reliable inferencing. In this paper, we presented four data gathering schemes that are applied candidates for WSNs. The energy efficiency of each scheme is analyzed to provide the basis for long-term, sustainable sensor operation in a WSN. According to the proposed data gathering schemes, the design of an intelligent sensor being capable of awareness of the potential energy consumption and automatically choosing the schemes for optimized energy efficiency is presented. Experimental results have shown that the structured WSNs achieve significant energy saving compared with the conventional centralized system and have potential for a wide range of applications in industry.

Acknowledgements

References

1 C. -Y. Chong and S. P. Kumar: Proc. IEEE. **91** (2003) 1247.
2 G. J. Pottie and W. J. Kaiser: Commun. ACM **43** (2000) 551.
3 T. Becker, M. Kluge, J. Schalk, K. Tiplady, C. Paget, U. Hilleringmann and T. Otterpohl: IEEE Sens. J. **9** (2009) 1589.
4 A. Tiwari, P. Ballal and F. L. Levis: ACM Trans. Sens. Netw. **3** (2007) 1.
5 F. Salvadori, M. de Campos, P. S. Sausen, R. F. de Camargo, C. Gehrke, C. Rech, M. A. Spohn and A. C. Oliveira: IEEE Trans. Instrum. Meas. **58** (2009) 3104.
6 D. Gallo, C. Landi and N. Pasquino: IEEE Trans. Instrum. Meas. **58** (2009) 3315.
7 R. Amirtharajah and A. P. Chandrakasan: IEEE J. Solid-State Circuits **33** (1998) 687.
8 C. B. Theurer, L. Zhang, D.O. Kazmer and R.X. Gao: IEEE Sens. J. **4** (2004) 28.
9 N. G. Elvin, N. Lajnef, and A. A. Elvin: Smart Mater. Struct. **15** (2006) 977.
10 T. Kurp, R. X. Gao, S. Sah: *Proc. IEEE Int. Instrum. Meas. Tech. Conf.*, Austin, Texas, May 3 −6, 2010, IEEE, Piscataway, New Jersey (2010) p. 93.
11 C. Alippi, G. Anastasi, M. Di Francesco and M. Roveri: IEEE Trans. Instrum. Meas. **59** (2010) 335.
12 A. Wang and A. Chandrakasan: IEEE Signal Proc. Mag. **19** (2002) 68.
13 R. X. Gao and Z. Fan: IEEE Trans. Instrum. Meas. **55** (2006) 415.
14 J. A. Sanchez, P. M. Ruiz, J. Liu and I. Stojmenovic: IEEE Sens. J. **7** (2007) 627.
15 C. Alippi, R. Camplani and M. Roveri: IEEE Trans. Instrum. Meas. **58** (2009) 3347.
16 P. Merlino and A. Abramo: IEEE Sens. J. **9** (2009) 1397.
17 C. -C. Shen, R. Kupershtok, S. Adl, S. S. Bhattacharyya, N. Goldsman and M. Peckerar: IEEE Sens. J. **8** (2008) 682.
18 S. Lindsey, C. Raghavendra and K. M. Sivalingam: IEEE Trans. Parallel Distrib. Syst. **13** (2002) 924.
19 M. X. Cheng and L. Yin: Int. J. Sens. Netw. **4** (2008) 48.
20 R. Yan, D. Ball, A. Deshmukh and R. X. Gao: *Proc. IEEE Sens. Conf.*, Vienna, Austria, October 24–27, 2004, IEEE, Piscataway, New Jersey (2004) p. 44.
21 Y. Xiang and V. Lesser: *Proc. 4th Int. Conf. on Multi-agent Syst.*, Boston, Massachusettes, July 10–12, 2000, IEEE, Piscataway, New Jersey (2000) p. 349.
22 C. Wang and R. X. Gao: IEEE Trans. Instrum. Meas. **49** (2000) 325.
23 R. Yan and R. X. Gao: Rob. Comput. Integr. Manuf. **21** (2005) 291.
24 Freescale Semiconductor: 56F8323/56F8123 Data Sheet, http://www.freescale.com (accessed on 2011).
25 RF Monolithics, Inc.: TR1001: 868.35 MHz Hybrid Transceiver, http://www.rfm.com (accessed on 2011).