Sensor placement strategy for high quality sensing in machine health monitoring

Robert X. Gao[†]

Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, MA 01003, USA

Changting Wang[‡]

Global Research Center, General Electric Corporation, Niskayuna, NY 12309, USA

Shuangwen Sheng[‡†]

Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, MA 01003, USA

Abstract.   This paper presents a systematic investigation of the effect of sensor location on the data quality and subsequently, on the effectiveness of machine health monitoring. Based on an analysis of the signal propagation process from the defect location to the sensor, numerical simulations using finite element modeling were conducted on a bearing test bed to determine the signal strength at several representative sensor locations. The results showed that placing sensors closely to the machine component being monitored is critical to achieving high signal-to-noise ratio, thus improving the data quality. Using millimeter-sized piezoceramic plates, the obtained results were evaluated experimentally. A comparison with a set of commercial vibration sensors verified the developed structural dynamics-based sensor placement strategy. It further demonstrated that the proposed shock wave-based sensing technique provided an effective alternative to vibration measurement, while requiring less space for sensor installation.

1. Introduction

Rotating machines and auxiliary components, such as spindles, gearboxes, and bearings, have been widely used in the industry and commercial fields (SKF 1996), (Baydar 2003). Due to adverse operation conditions, improper installation, and material fatigue, the health status of a rotating machine will deteriorate over time, leading to machine failures. Premature and unexpected machine failures (e.g. due to defective bearings) may not only affect product quality, cause significant capital loss, but even endanger human lives (Tepfer and Baldor 1996). Thus, effective and efficient in-situ

---

†Professor, E-mail: gao@ecs.umass.edu
‡NDE Scientist, E-mail: wangch@crd.ge.com
‡†Research Assistant, E-mail: ssheng@ecs.umass.edu

bearing condition monitoring and fault diagnosis is highly desired to capture the defect at its incipient stage.

As a basic element supporting rotating machine systems, rolling bearings have been employed in virtually all branches of industrial and commercial applications. One of the typical bearing failure modes is localized spalling on the contact surfaces of bearing elements (Mori, *et al.* 1996), (Mano, *et al.* 2000), (Nelias 1999). When rolling elements roll over the defect, impulsive impacts are generated, causing the bearing structure to vibrate. Accordingly, the intensity of bearing vibrations can be used as a representative indicator for monitoring the bearing operating conditions. However, due to vibrations of the machine structure and background noise often encountered in a machinery environment, signals generated by a defective bearing may be contaminated, and poor signal-to-noise ratio (SNR) makes it difficult for effective signal identification and analysis (Gao and Phalakshan 1995). The difficulty will be compounded if the locations of the condition-monitoring sensors are not properly chosen. While sophisticated signal filtering and analysis techniques offer certain help, such approaches generally come with an increased computational load, thus hindering their real-world applications on the factory floor, especially when real-time analyses are required for in-situ machine condition monitoring. Research is therefore necessary for developing sensor placement strategies to optimize defect detection mechanisms and enable high quality sensing (Holm-Hansen and Gao 1997, 2000).

This paper investigates the effect of structural dynamics on sensor placement and subsequently, on the quality of the measured data, which directly affects the effectiveness of machine health monitoring. Theoretical consideration on the mechanisms of signal attenuation from the source of generation to the sensor location is first introduced. Numerical simulation using the finite element modeling (FEM) is then conducted to calculate the signal strength at several representative sensor locations on a rolling bearing as an application example. The simulation provides a quantitative justification on why placing sensors closely to the machine component being monitored is of critical importance to ensure high quality sensing. The simulated result is then experimentally verified by comparing customized sensors placed closely to a test bearing with commercial sensors placed away from the bearing. To enable the design of miniaturized sensors for bearing structural integration, issues involving sensing mechanism, effect of proof mass, mass loading on sensor sensitivity, and trade-off between sensitivity and effective bandwidth of the sensor are studied. It was found that vibration sensing based on compressive shock wave instead of on the traditionally acceleration-based technique provides a good alternative for high quality data acquisition. Such a configuration reduces the need for space to accommodate the proof mass, thus minimizing the space requirement for sensor installation. The developed sensing approach, integrated with advanced signal processing techniques such as wavelet and wavelet packet analysis (Yan and Gao 2004), (Wang and Gao 2003), (Ikolaou and Ntoniadis 2002), neural-fuzzy data fusion (Huang, *et al.* 2002), (Sheng and Gao 2003), and complexity measures (Yan and Gao 2004), has the potential to realize a "diagnosis-on-a-chip" that is not only applicable to bearing condition monitoring, but also the health diagnosis of a wide range of mechanical systems.

This paper is organized as follows: section 2 discusses the mechanism of structural attenuation when a signal propagates from the source of initiation to the sensor, followed by a finite element simulation of optimal sensor placement. Issues regarding the design of integrated sensor and shock wave-based sensing mechanism are then discussed in section 3. Also discussed are the effects of proof mass, mass loading, and effective bandwidth. Section 4 presents the experimental results obtained from customized and commercial sensors, whereas conclusions are drawn in section 5.

2. Signal attenuation mechanisms

To measure vibration signals generated by a rolling bearing, accelerometers have been commonly installed on the bearing fixture or adjacent machine housing. Since bearings are mechanically coupled with other machine structures (e.g. spindles, gearboxes, or motors), accelerometers mounted on the machine housing will pick up structural-borne vibrations and environmental noise, in addition to the bearing defect impact-induced vibrations. The result is a low signal-to-noise ratio (SNR). This is particularly noticeable at the defect incipient stage, when the defect-induced impact is weak in magnitude and of short duration, because of the small geometrical size of the defect. Thus, understanding the mechanism of structural attenuation of a signal during its propagation is key to devising a sensor placement strategy that optimizes its effectiveness.

2.1. Structural attenuation

Assuming a point source where transient mechanical waves due to localized structural defects inside of a bearing (e.g. on the outer raceway, inner raceway, or rolling elements) are generated, as illustrated on the left hand side in Fig. 1. The amplitude (*H*) of the waves propagating through the bearing can be expressed as (Kolshy 1963, Degtyar 1998):

$$H(r, t) = \frac{A}{r} e^{-j\phi(r, t)} \tag{1}$$

where *r* is the distance away from the source, *A* is an amplitude constant depending on the source vibration strength, and ϕ represents the phase function of distance *r* and time *t*. This formulation indicates that the amplitude of the measured vibration signals is attenuated proportionally to the signal traveling distance.

Another source of signal attenuation is caused by the structural damping. Shown on the right hand side in Fig. 1 is a realistic bearing impulse response obtained from a bearing test bed (Fig. 2). The signal amplitude decreases quickly with time, indicating the existence of large structural damping. Such damping is commonly seen in machine systems employed in real world applications. As an example, Fig. 2 illustrates the cross-sectional view of a bearing-spindle test bed where a defect-induced impact was originated at the outer raceway of the test bearing. The vibration signals will propagate through the rolling elements, the inner raceway, the shaft, the bearing housing, the support bearing, and other structural
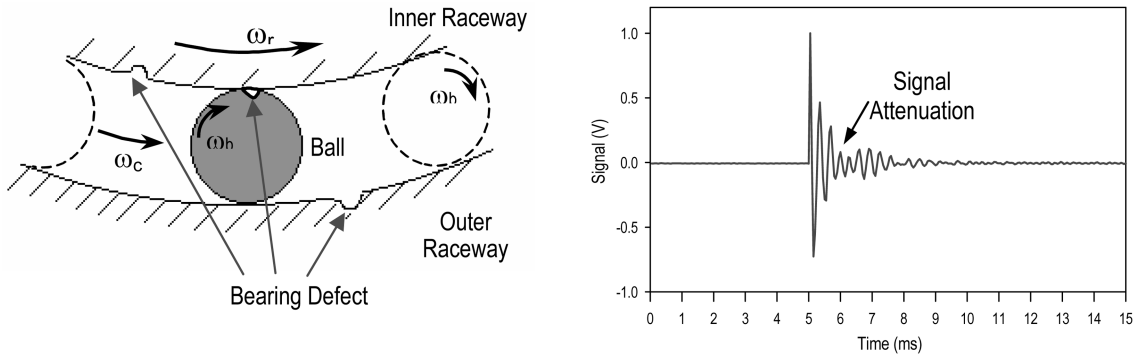
Fig. 1 Bearing defect-induced signal generation (left) and structure response (right)
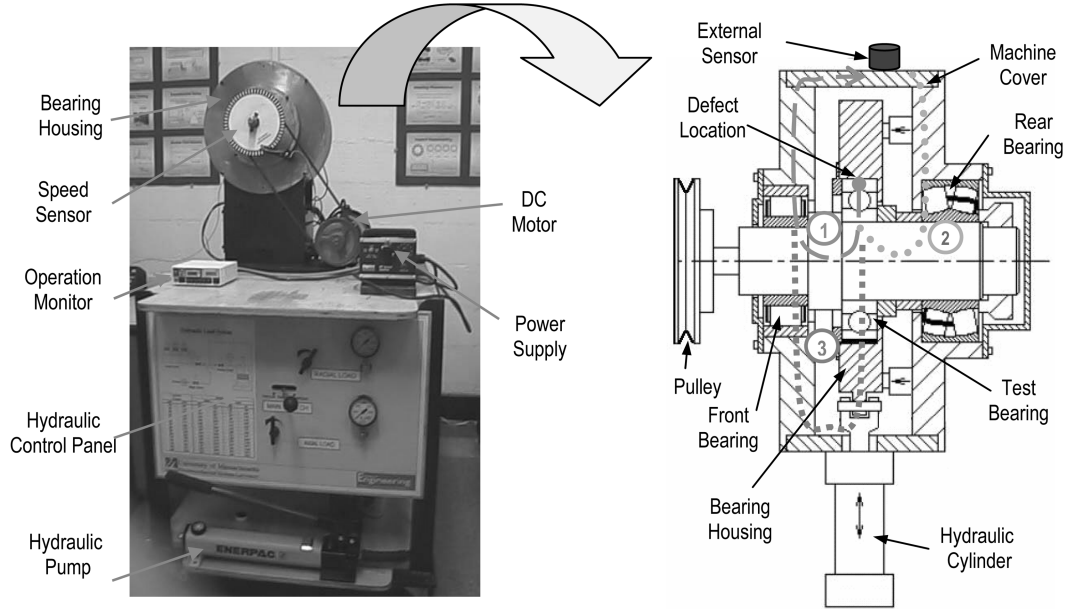
Fig. 2 A bearing-spindle test bed (left) and its cross-sectional view (right)

components, before reaching a vibration sensor placed on top of the bearing housing. Attenuation of a defect-related vibration signal as it propagates through such a long chain of components can be significant. In reality, the signal may propagate along multiple paths (e.g. paths ① through ③), before reaching the sensor. As a result, the energy contained within the signal will be dissipated at multiple locations, making it difficult for a sensor to detect the existence of the defect-related signal.

Consider the signal propagation path #1 illustrated in Fig. 3. The path is modeled as a chain of spring-mass-damper sub-systems (m_i and k_i, $i=1,2,\ldots j$), connected in series to each other. Each sub-system along the propagation chain is characterized by a damping ratio (ξ_i), which dissipates energy contained within a signal, leading to signal attenuation. For purpose of simplicity, a sinusoidal signal wave propagating through the chain is assumed. The energy dissipated due to structural damping can be calculated from the work performed by the damping forces f_{di} along the displacement u_i of the mass element m_i, as given by:
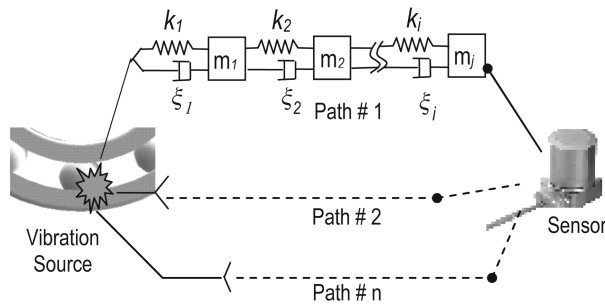

Fig. 3 Signal propagation paths from a point source to the sensor location

$$E_{d_i} = \int_T f_{d_i} du_i \tag{2}$$

where T is the duration of the transient vibration (the sinusoidal wave). Let c_i be the damping factor corresponding to the damping ratio ξ_i, and substitute $f_{d_i} = c_i \dot{u}_i$ and $u_i = A_i$ and $\left(\frac{2\pi}{T}t\right)$ into Eq. (2). With $d_{u_i} = \dot{u}_i dt$, it can be derived that the energy dissipation at the i-th sub-system along the signal propagation chain is given by:

$$E_{d_i} = 2\pi\xi_i \frac{T_{n_i}}{T} k_i A_i^2 \tag{3}$$

where T_{n_i} is the natural period of the vibration system. This indicates that the energy loss in each spring-mass-damper sub-system is proportional to the square of the signal amplitude and inversely proportional to the duration of the transient vibration. This result agrees with the observation that higher frequency vibration components generally suffer more attenuation than lower frequency components, due to the low-pass characteristic of mechanical systems. Since transient vibrations caused by defect-related impacts are usually of short duration, because of the small sizes of defect at the incipient stage, such vibration will be considerably attenuated along the signal propagation path. This makes it difficult for sensors placed at a distance away from the defect location to accurately and timely detect the defect signals.

Based on Eq. (3) for damping-induced energy dissipation, an energy loss ratio can be defined, which is the ratio between the actual energy content available at the i-th sub-system and that under an ideal situation, when no energy loss exists. The latter can be expressed as $\frac{1}{2}k_i A_i^2$, and subsequently, the energy loss ratio is given as:

$$L_i = 4\pi\xi_i \frac{T_{n_i}}{T} \tag{4}$$

Let E_{p1} be the energy at the defect impact location ("Vibration Source" in Fig. 3), and the vibration signal is assumed to propagate through path #1. The energy transmitted to the location of the external sensor through path #1 can be calculated as:

$$E_{o1} = E_{p1} \prod_{i=1}^{j} (1 - L_i) \tag{5}$$

The vibration magnitude A_j at the sensor location can be determined by letting $E_{o1} = \frac{1}{2}k_i A_j^2$. To illustrate the energy loss due to structural damping, consider ten identical spring-mass-damper sub-systems that are connected in series, with a damping ratio of $\xi_i = 0.02$ and vibration period $T_{n_i} = T$. In each sub-system, the energy loss ratio L_i is calculated to be 25%, using Eq. (4). The total energy loss at the end of the propagation chain where the sensor is located will be 95%, and the signal amplitude will be attenuated by more than 75%.

Non-elastic materials such as lubricating grease also contribute to signal attenuation along its propagation path. If dissipative materials are included within the path, the vibration energy loss will express itself as heat generation. Furthermore, boundary conditions of the media also have a significant effect on the wave propagation. For example, signal reflection occurs at interfaces among dissimilar materials, even when sufficient pressure was applied at the interface (e.g. through tight interference of the machine
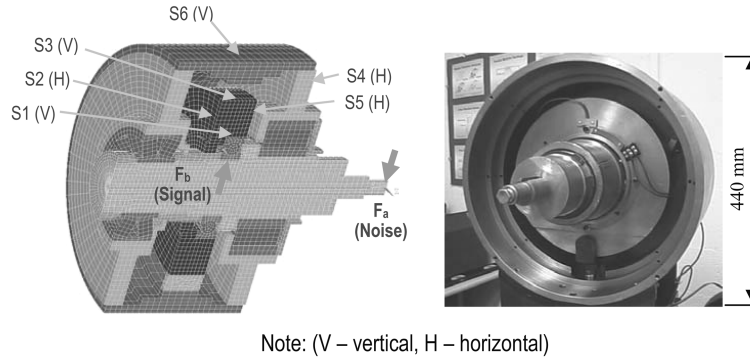
Note: (V – vertical, H – horizontal)

Fig. 4 Finite element model (left) of the bearing test bed (right)

assembly, Rinehart 1975, Degtyar 1998). Given a normal incident wave with amplitude H_1 propagates through a cohesive interface between two dissimilar materials, the energy reflection ratio of source wave H_1 is:

$$\gamma_r = 1 - \frac{H_2}{H_1} = 1 - \frac{4\rho_1 c_1 \rho_2 c_2}{(\rho_1 c_1 + \rho_2 c_2)^2} \tag{6}$$

where H_2 is the transmitted wave, ρ_1, and ρ_2 are material densities and c_1, c_2 the wave propagation speeds of the two materials. As an example, when vibration signals propagate through a steel-aluminum interface, the reflection ratio can be as high as 20%, if ρ_1=8,030 kg/m³, ρ_2=2,770 kg/m³, c_1=5,790 m/s, and c_2=6,420 m/s.

The above analysis demonstrates that, to retain the maximal energy content in the signal, which is a measure for the signal's "detectability" under given operational conditions, condition monitoring sensors shall be placed as closely as possible to the machine component or structure to be monitored. A shorter distance between the sensor and the defect location helps achieve a higher SNR. Also, it reduces exposure of the signal to background contaminations (e.g. vibration from other machine structures, signal interference). Because of the complexity of realistic machine systems, parameters illustrated in the model in Fig. 3 may not be readily available, and accurately calculating the energy loss ratio L_i can be thus difficult. To provide a reference framework for establishing a sensor placement strategy for machine condition monitoring, numerical analysis using FEM was performed, based on the geometry of the bearing test bed shown in Fig. 4.

Table 1 Material constants used for the modeling

Material	Steel AISI-304	Steel AISI-C1020	Aluminum 2024-T3
Young's Modulus (Pa)	1.93e11	2.07e11	7.30e10
Poisson's Ratio	0.28	0.29	0.33
Density (g/cm³)	8.03	7.85	2.77
Damping Factor	0.01	0.01	0.01
Parts in the 3D Model	Left support Right support Housing Shaft base	Bearing 1 Bearing 2 Bearing 3	Top cover

Table 2 Relative signal strength at specified sensor locations

	Y1	Y2/Y1	Y3/Y1	Y4/Y1	Y5/Y1	Y6/Y1
Signal from F_b	S_1	49.5%	78.5%	20.6%	57.9%	27.4%
Noise from F_a	N_1	13.9	1.4	11.6	14.8	2.9
S/N (dB)	S_1/N_1	-29	-5	-35	-28	-20

2.2. Structural response

A three-dimensional finite element model of the bearing test bed was established based on the material properties given in Table 1. To investigate the effect of sensor placement on signal strength in response to an impulsive input and subsequently, on the sensing quality, a total of six representative sensor locations, noted as S1 through S6, were identified, as shown in Fig. 4.

Selection of the representative sensor locations was based on two aspects: 1) common measurement practice and accessibility to the test bed, and 2) realistic locations where sensors will be placed in experimental studies. Boundary conditions used in the finite element (FE) modeling were defined to match the realistic test bed system. For example, when setting up the FE model for the bearing, contact elements were used to reflect the interaction between the rollers and the bearing raceways. Similarly, to model the bearing housing structure, the degreeof-freedom along the z-axis was constrained such that only responses in the *x* and *y* planes were considered. The direction of the applied signal and noise input (F_a and F_b) was chosen to match realistic experimental scenarios in the bearing test bed system, where vibration input along the vertical direction has shown to produce the strongest structural response.

To simulate the impulsive impact from defect-induced vibrations to the bearing structure, a dynamic input (Signal, F_b) of 16,000 N and 3 ms duration was applied to the test bearing, which is located in the center of the test bed. Furthermore, to simulate machine-borne vibrations from other structural sources, an input of the same magnitude and duration (Noise, F_a) was applied to the end of the driving shaft. Transient response of the structure to the impulse forces was analyzed, using the software package ANSYS. In Table 2, the relative signal strength at the specified sensor locations (based on the acceleration values) are illustrated.

As shown in Table 2, the structural vibration amplitude at sensor location S_1, in response to signal input F_b, is referred to as S_1, whereas response to the noise input F_a is defined as N_1. The pair S_1-N_1 serves as the basis for comparisons between the signal strength at location S_1 and the rest five locations. Under the impact of F_b, a vibration sensor (e.g. accelerometer) placed at location S6 (bearing housing cover) sensed only 27.4% of the signal magnitude, as compared to a sensor placed on location S_1. This indicates that a sensor placed near the bearing (i.e. source of signal generation) is more sensitive in capturing the effect of the structural defect than sensors located farther away from the bearing. On the other hand, for structure-borne vibrations as simulated by the impulsive input (F_a) on the shaft, the signal magnitude measured by an accelerometer at location S1 was about 1.4 to 14.8 times weaker than the same sensor placed on locations S2-S6. This means that a sensor placed closer to the bearing would be less disturbed by structure-borne vibrations, and therefore has a higher immunity against background noise. Using the SNR at location S1 as the reference base, it is seen that the ratios at other locations are about 5 to 35 dB lower. As a result, signals induced by the bearing defects will only reach these sensors after strong or very strong attenuations, which may render the defect sensing completely

ineffectual. While the specific machine structure may vary in each application, the above simulation based on a realistic bearing test bed illustrates a generally applicable principle that, to achieve high quality sensing, the monitoring sensors should be placed as closely as possible to the machine component or structure to be monitored.

3. Sensor design

To experimentally evaluate the sensor placement strategy as discussed above, three customized sensors, one for dynamic force measurement and two for vibration measurement, were developed. Critical sensor design aspects, e.g., sensor characteristic, sensitivity, and frequency response, were investigated.

3.1. Sensor characteristic

A sensor module was designed to fit in a slot of 4.5×18×30 mm in the outer raceway of a type 6220 bearing. As illustrated in Fig. 5, the module consisted of a piezoceramic plate for dynamic force measurement, signal conditioning electronics, and a wireless transmitter, all packaged within a rectangular housing. One important parameter considered in the sensor design was the resonant frequency of the piezoceramic plate. The lowest resonant frequency of a piezoceramic plate f_r is determined by the frequency constant k_f and the thickness d of the plate (Randeraat and Setterington 1974) as:

$$f_r = k_f/d \qquad (7)$$

Thus, a piezoceramic plate with a frequency constant $k_f = 2,032$ kHz-mm and thickness $d = 2.5$ mm would have the lowest resonance frequency being about 800 kHz. Reducing sensor thickness will increase the sensor resonant frequency and subsequently, its bandwidth, but make the sensor more susceptible to dielectric breakthrough due to high voltage, and to bending breakage due to imperfect

1. Electronic components   2.Circuit board   3. Piezoceramic plate
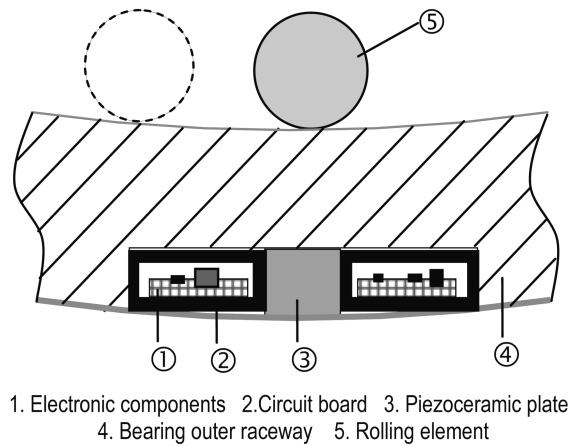4. Bearing outer raceway   5. Rolling element

Fig. 5 Embedded dynamic force sensor module

installation surface.

Upon application of a dynamic force, the piezoceramic sensor generates an electrical charge output Q, which is converted to an output voltage V_o by a charge amplifier. Derivations yielded the relationship between the applied dynamic force *F* and the voltage output of the charge amplifier as:

$$V_o(s)/F(s) = \varepsilon^{-1} \cdot \frac{Rs}{1 + RCs} \tag{8}$$

The time constant of $\tau = RC$ determines the 3 dB cut-off frequency. For a circuitry with $R = 5$ MΩ and $C = 33$ nF, $1/(2\pi RC) \approx 1$ Hz. This indicates that the circuit will be able to measure bearing signals down to several Hz, without causing signal distortion. Increasing the feedback capacitance *C* of the charge amplifier will improve the circuit's low frequency response, but reduce the sensitivity of the sensor.

In addition to the dynamic force sensor configuration, two shock wave sensors were designed using piezoceramic plates. Compared to an accelerometer that has a proof mass attached to the sensing element, the shock wave sensors did not contain a proof mass, and hence reduces the space required to accommodate the sensor. Such a design is advantageous in view of sensor structural integration within the host structure. The shock wave sensors were glued to the bearing housing. Impulsive impact from the rolling ball-raceway defect interactions generates shock waves that compress the sensors and induce a microscopic deformation within the sensor, in the form of mechanical strain ($\delta l/l$). Because of the strain development, an electrical charge will be produced that is directly proportional to the magnitude of the compressive wave. Such a charge generation mechanism is different from that of an accelerometer, where a proof mass physically compresses the piezoceramic plate placed beneath it, when subject to an accelerated motion.

To understand the difference between the shock wave sensors and the accelerometers, a piezoceramic plate with a proof mass (*M*) placed on top of it was modeled as an equivalent mass-spring-damper system. Let m_e and *k* be the equivalent mass and stiffness of the piezoceramic plate, respectively. When subject to an acceleration input $a*\sin(\omega t)$, the proof mass *M* has a displacement response determined by (Inman 1996):

$$y = \frac{1}{\omega_n^2} \frac{a}{\sqrt{[1 - (\omega/\omega_n)^2]^2 + 4\xi^2(\omega/\omega_n)^2}} \tag{9}$$

where ξ is the damping ratio and ω_n is the natural frequency of the accelerometer:

$$\omega_n = \sqrt{\frac{k}{M + m_e}} \tag{10}$$

For proper functioning of the accelerometer, ω_n should be much higher than the frequency ω of the vibration to be measured. Thus for $\omega/\omega_n \ll 1$, the response *y* can be considered as independent of ω, and Eq. (9) is simplified as:

$$y = \frac{a}{\omega_n^2} = \frac{M + m_e}{k}a \tag{11}$$

The displacement *y* of the proof mass causes the piezoceramic plate to be compressed, which in turn generates an electric charge (*Q*). Since *y* is proportional to the acceleration *a*, as indicated in Eq. (11), the charge output will be proportional to *a* as well. Thus, increasing the mass *M* and/or m_e will increase the accelerometer's sensitivity linearly. At the same time, since the bandwidth of an
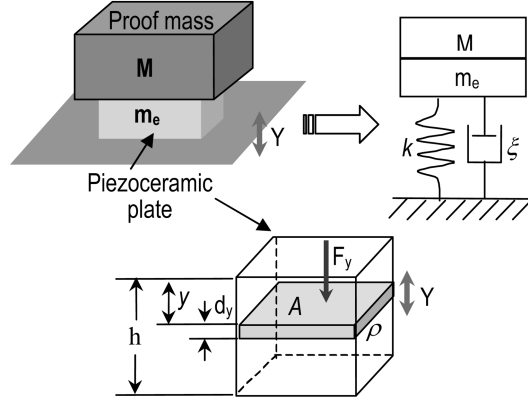
Fig. 6 Effect of proof mass on the output characteristic of shock wave sensor

accelerometer is inversely proportional to the sum of M and m_e, as illustrated in Eq. (12), an increase in the mass M and/or m_e will result in a decrease of the bandwidth (BW_{3dB}), which can be estimated to be 2/3 of ω_n (Randeraat and Setterington 1974):

$$BW_{3dB} = \frac{2}{3}\omega_n = \frac{2}{3}\sqrt{\frac{k}{M + m_e}} \tag{12}$$

When the proof mass is removed to minimize the sensor spatial requirement, no external mass loading will be exerted on the piezoceramic plate upon acceleration (vibration). The relationship between the charge output of the sensor and the input acceleration (i.e. the transfer function of the sensor) can then be obtained by considering the configuration in Fig. 6, where h, A and ρ represent the thickness, cross sectional area, and density of the piezoceramic plate, respectively.

Consider an infinitely thin slice within the piezoceramic plate, dy, at the coordinate y relative to the reference plane. When the plate is subject to an acceleration a, the slice will be compressed by the inertial force F_y, which is given by $F_y=[(h-y)A\rho]a$, where the expression $[(h-y)A\rho]$ accounts for the mass of the portion on top of the slice. The piezoelectric voltage sensitivity g_{33} for the slice defines the electric field per unit stress (Randeraat and Setterington 1974), and is expressed as:

$$g_{33} = \frac{electric\ field\ within\ slice}{unit\ stress\ inside\ slice} = \frac{dV/dy}{F_y/A} \tag{13}$$

where dV is the voltage on the slice that is generated under the force F_y, and dy is the thickness of the slice. Therefore,

$$dV = \frac{dy}{A}g_{33}F_y = g_{33}(h-y)\rho a dy$$

The total voltage output from the chip is the integration of dV over the thickness:

$$V = \int_{y=0}^{h} g_{33}(h-y)\rho a dy = \frac{1}{2}g_{33}\rho a h^2 \tag{15}$$

The charge Q generated on the plate is the voltage V multiplied by the plate capacitance C, which is given as $C=\varepsilon A/h$, with ε being the dielectric constant of the piezoceramic plate. As a result,

$$Q = \frac{1}{2}g_{33}\varepsilon A \rho h a = \frac{1}{2}g_{33}\varepsilon\, m_p a \tag{16}$$

where $m_p = A\rho h$ is the mass of the plate. From Eq. (16), it is evident that charge generation in a shock wave sensor is proportional to the mass of the piezoceramic plate, as well as to the acceleration that the plate is subject to. Thus, with the material constants (g_{33} and ε) and mass of the piezoceramic plate (m_p) known, the acceleration a can be measured directly from the charge output as: $a = 2\dfrac{Q}{(g_{33}\varepsilon) \times m_p}$. In comparison, a conventional accelerometer measures acceleration through the relationship $a = \dfrac{Q}{(g_{33}\varepsilon) \times m}$, with m representing the combined mass of the piezoelectric sensing element (e.g. a piezoceramic plate) and the proof mass.

3.2. Sensor sensitivity

To quantitatively investigate the effect of proof mass and shock wave-based vibration sensing, a series of proof masses were manufactured. These proof masses were placed on top of the piezoceramic plate, individually and in combinations, to realize a mass ratio (proof mass-toplate mass, or m/m_p) of up to 34. The proof mass-piezo plate pair was subsequently mounted on top of an electromagnetic vibration shaker, which outputs oscillations within the frequency range of 25-5,000 Hz. By observing the signal output of the piezoceramic plate under various proof mass ratios, the sensitivity and bandwidth of the accelerometer were obtained. To provide a reference base for the measurement, a commercially available Kistler accelerometer (model 8712A5M1) was used, which has a sensitivity of 992 mV/G, with $\pm$5G accuracy.

As shown in Table 3, the sensitivity of the sensor increases as the mass ratio increases (i.e. when

Table 3 Sensor sensitivity at 100 Hz

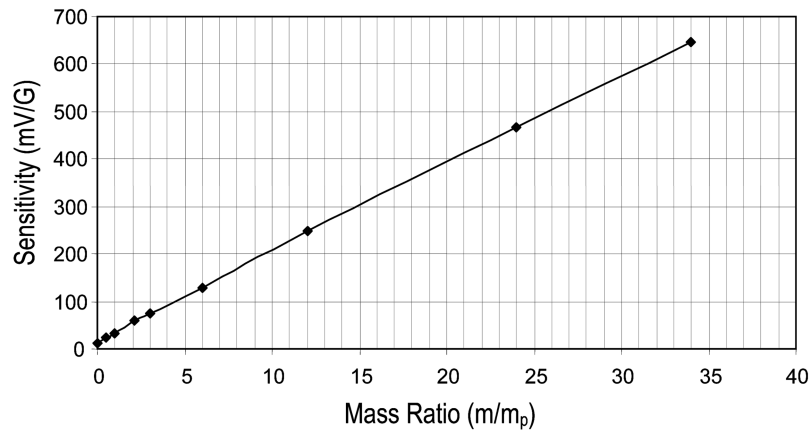| Mass Ratio (m/m_p)	0	0.5	1	2	3	6	12	24	34
Sensitivity (mV/G)	11	25	38	65	88	144	247	466	646

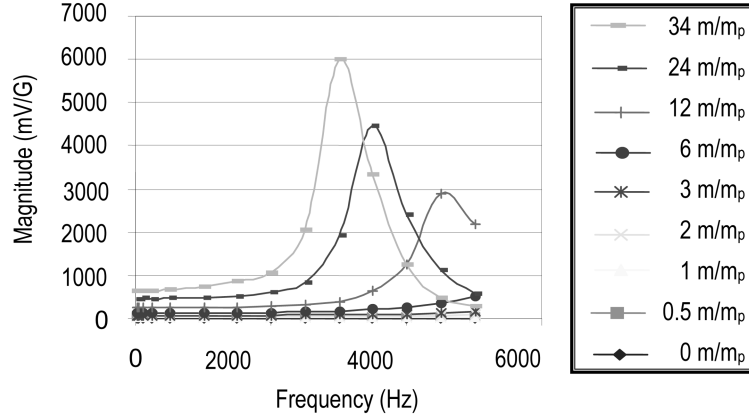Fig. 7 Relationship between the mass ratio and sensor sensitivity

Fig. 8 Sensor frequency response vs. mass ratio

larger proof mass was placed on top of the piezoceramic plate). The relationship is linear, as illustrated in Fig. 7. The sensitivity values were obtained under a vibration shaker frequency of 100 Hz, which is the standard frequency commonly used for calibrating the sensitivity of commercial accelerometers.

When no proof mass was placed on the piezoceramic plate, the sensitivity of the sensor was measured as 11 mV/G. This situation can be modeled as an accelerometer consisting of a massless piezoelectric sensing element and an equivalent seismic mass m_e that is half of the sensor's total mass, i.e., 0.4 g for a piezoceramic plate of 0.8 g. Based on the previous discussion and an amplification factor of 3.1 V/N for the signal conditioning electronics, the 0.4 g equivalent mass would apply an inertial force of F=3.9 mN to the assumed massless sensing element, under 1G acceleration. Under this inertial force, a voltage output of V=12.1 mV would be generated, which agreed with the experimentally determined sensitivity of 11 mV/G within 9%.

3.3. Sensor frequency response

While increasing the proof mass would increase the sensor sensitivity, it has the drawback of reducing the sensor's effective bandwidth. This is expressed by the relationship between the natural frequency ω_n of the sensor system (essentially a mass-spring-damper system) and its effective mass, as indicated in Eq. (10). In Fig. 8, the frequency response of the piezoceramic chip under various mass ratios is shown. Under low mass ratios (e.g. < 3 m/m_p), the response curve remained flat up to about 5 kHz. As the mass ratio increases, the flat curve region becomes narrower. At 24 m/m_p, the region is reduced to less than 2 kHz. The flat region represents the effective bandwidth of an accelerometer, within which measurements are made. Thus, a shock wave-based sensor has in general a wider bandwidth than an accelerometer, making it more flexible for vibration measurements in a broader range.

Given the complex effect of the proof mass on the sensitivity and effective bandwidth of the sensor, selection of a specific proof mass value should be determined in accordance with each specific bearing monitoring task. Two factors that need to be considered are 1) the frequency range of the input signal, and 2) the spatial constraint to which the sensor is subject. As an example, the existence of structural defect within a bearing may be identified by the repetitive "roll-over" frequency of the measured signal. Such a frequency is typically associated with the periodic interactions between the rolling elements of

the bearing and the localized defect (on the surface of bearing raceway or the rolling elements). Every time when a ball rolls over the defect, an impulsive impact will be generated. Given the repetitive nature, such roll-over frequencies are of relatively low frequency, e.g., under 1 kHz. In such cases, a relatively large proof mass (e.g. mass ratio of 15-20) can be used to maximize the sensitivity of the sensor, if sufficient space is available to accommodate it. On the other hand, if higher signal frequency need to be detected, e.g., when the bearing rotational speed is higher, or instead of the repetitive roll-over frequency, the characteristic frequency associated with the defect inception and propagation at the microscopic level is of interest to the monitoring task (where the frequency range of interest may extend beyond 20 kHz or higher), then only small proof mass or no proof mass can be considered to provide the needed bandwidth, although with the drawback of a lower sensitivity. In this context, miniaturization technology such as MEMS (micro-electro-mechanical systems) may provide significant advantage. By minimizing the dimension and consequently, the mass of the sensing element itself, the ratio of the proof mass to the sensor's mass can be increased significantly to increase the sensitivity of the sensor, while still maintaining a minimal sensor package.

4. Experimental results

The small size of the piezoceramic plates (in the millimeter range) made it feasible for them to be placed closely or directly into the bearing test bed (Fig. 2) to verify results from the finite element analysis on how sensor locations affect measurement data quality. Additionally, experiments were conducted to investigate various sensing mechanisms for effective bearing defect detection. The results were benched through "side-
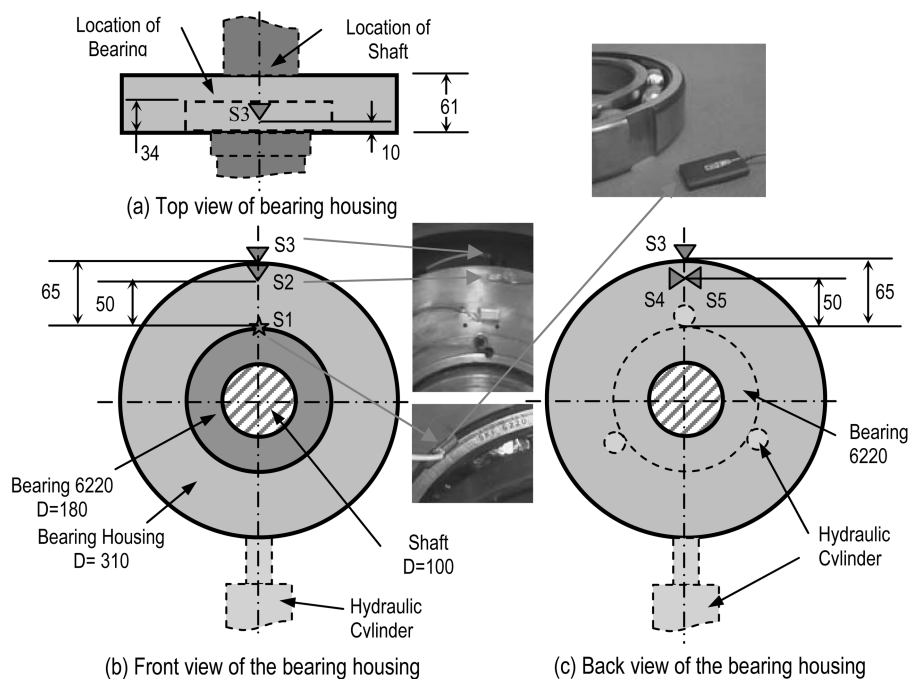

Fig. 9 Sensor locations on a bearing test bed (Unit: *mm*)

Table 4 Experimental sensor specifications

Sensor notation	Sensor type	Physical quantity measured	Sensor location
S1	Piezoceramic plate (UMass)	Dynamic force	Within bearing outer raceway
S2	Piezoceramic plate (UMass)	Shock wave/strain	Bearing housing front
S3	8636B50 (Kistler)	Acceleration	Bearing housing top
S4	Piezoceramic plate (UMass)	Shock wave/strain	Bearing housing back
S5	8636B50 (Kistler)	Acceleration	Bearing housing back
S6	CMSS787A (SKF)	Acceleration	Top of the bearing test bed cover

by-side" comparison between the custom-designed sensors and three commercial sensors.

Fig. 9 illustrates the investigated sensor locations in the experiment. A total of six representative sensor locations (S1 through S6) were selected, with S6 being on the top of the bearing test bed cover (not shown in the Figure).

Selection of the representative sensor locations was based on common measurement practice and accessibility to the test bed. The three sensors at S1, S2, and S4 consist of the same piezoceramic plate (PKI502 material) with a piezoelectric constant (g_{33}) of 400 pC/N and a dimension of 10×4.5×2.5 mm. The sensor at S1 was configured as a dynamic force sensor and embedded into the cavity of a phenolic module, which in turn was sandwiched between the bearing outer raceway and the bearing housing. The signal conditioning circuitry had a sensitivity of 2.5 mV/N. The specific physical quantity measured by each individual sensor is listed in Table 4.

The signals measured at the sensing locations S1, S2, and S4 under 1,000 rpm and 11 kN radial load are shown in Fig. 10. The spectrum of each signal is given to the right of the time domain presentation. The prefix "*a*", "*b*" and "*c*" in the frequency labels specify components from the test bearing (6220) and two support bearings (2322 and 22318), respectively.

The sensor at S1 measures dynamic load variations due to the balls rolling over the slot on the raceway, where the sensor is placed. As shown in Fig. 10 (top), vibrations caused by bearing unbalance and misalignment were measured at 1,000 rpm (equal to f_u=16.7 Hz) and 2,000 rpm (f_m=33.3 Hz), respectively. Strong signal components can be seen at the Ball Passing Frequency in Outer-raceway (BPFO) and its harmonics. This frequency is the same as the repetitive frequency of the defect signals, caused by the ball-defect interactions. Thus, strength and variation of the BPFO frequency is a measure for the existence and propagation of bearing structural defects.

Given that the sensors located in S1 and S2 measure two different physical parameters initially (dynamic force variation and shock wave-induced strain), a direct comparison of their measurement results is not feasible. To provide a general overview as to which type of sensing mechanism enables better identification of the bearing frequency component of interest (BPFO), the sensors output was normalized against the peak values of their respective BPFO frequency.

As shown in Fig. 11, both sensors have been able to identify the BPFO components effectively. However, compared to S1, the S2 sensor was more effective in measuring the unbalance (f_u) and misalignment (f_m) frequencies contained in the bearing signals. This indicates that the S2 sensor is better suited for detecting structural unbalance and misalignment.

In addition to more effectively detecting unbalance and misalignment, the S2 spectrum indicates a frequency component at the Ball Passing Frequency Inner-raceway (BPFI), which is reflective of a structural defect in the bearing inner raceway. Comparison with the spectrum of a "healthy" bearing
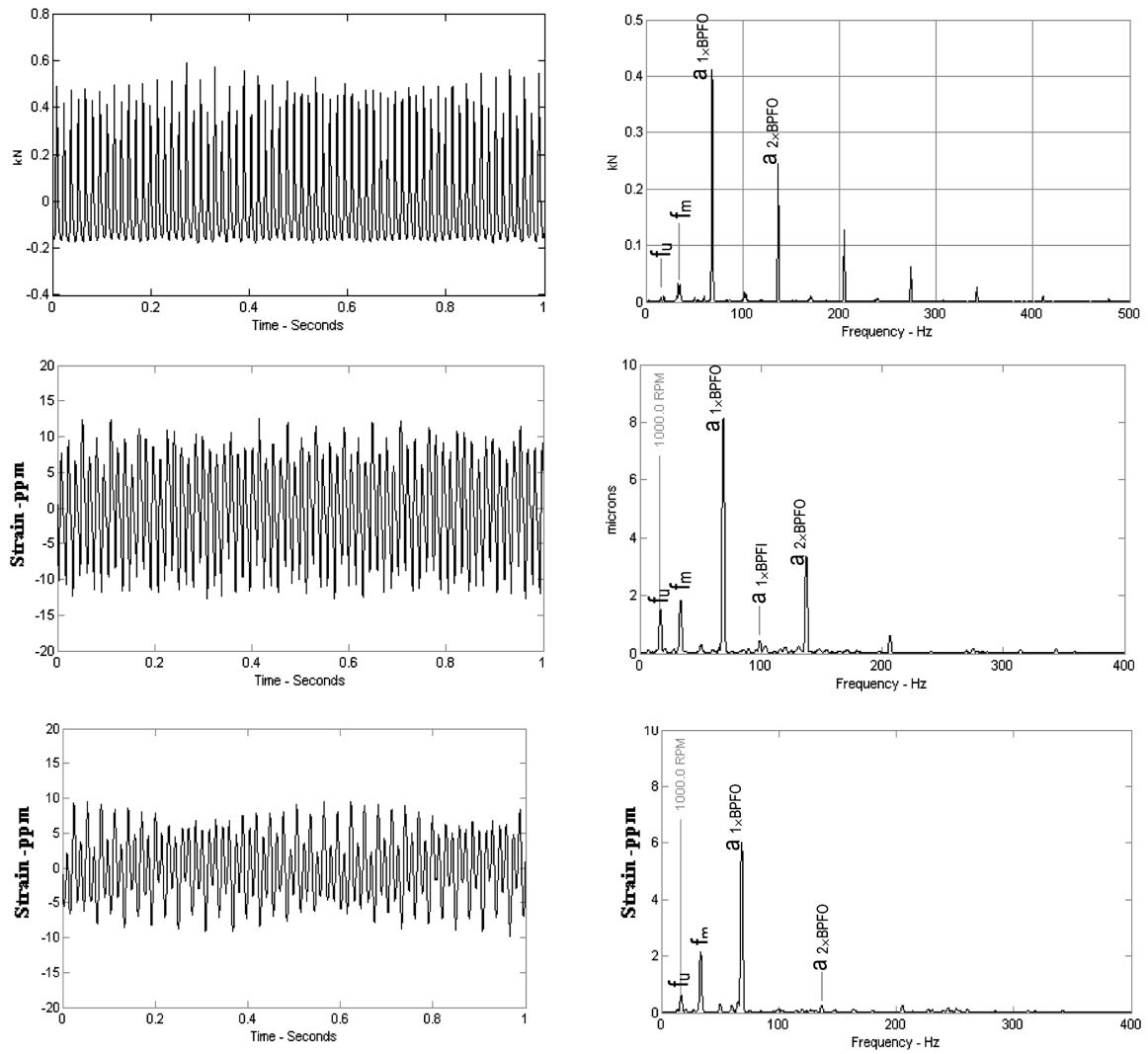
Fig. 10 Signals measured at sensor locations: S1 (top), S2 (middle), and S4 (bottom)
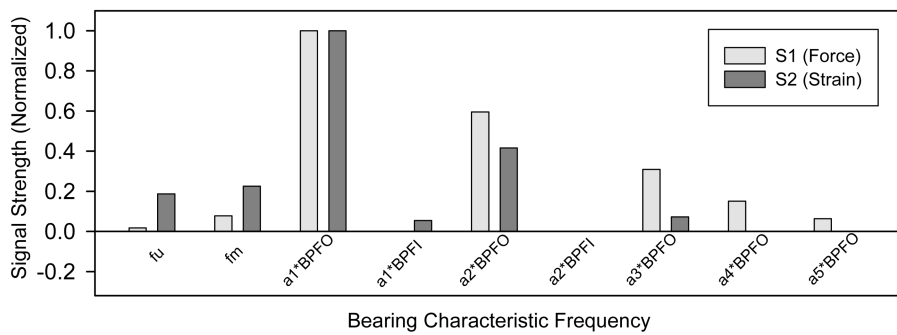

Fig. 11 Normalized signal strength at sensor locations S1 and S2
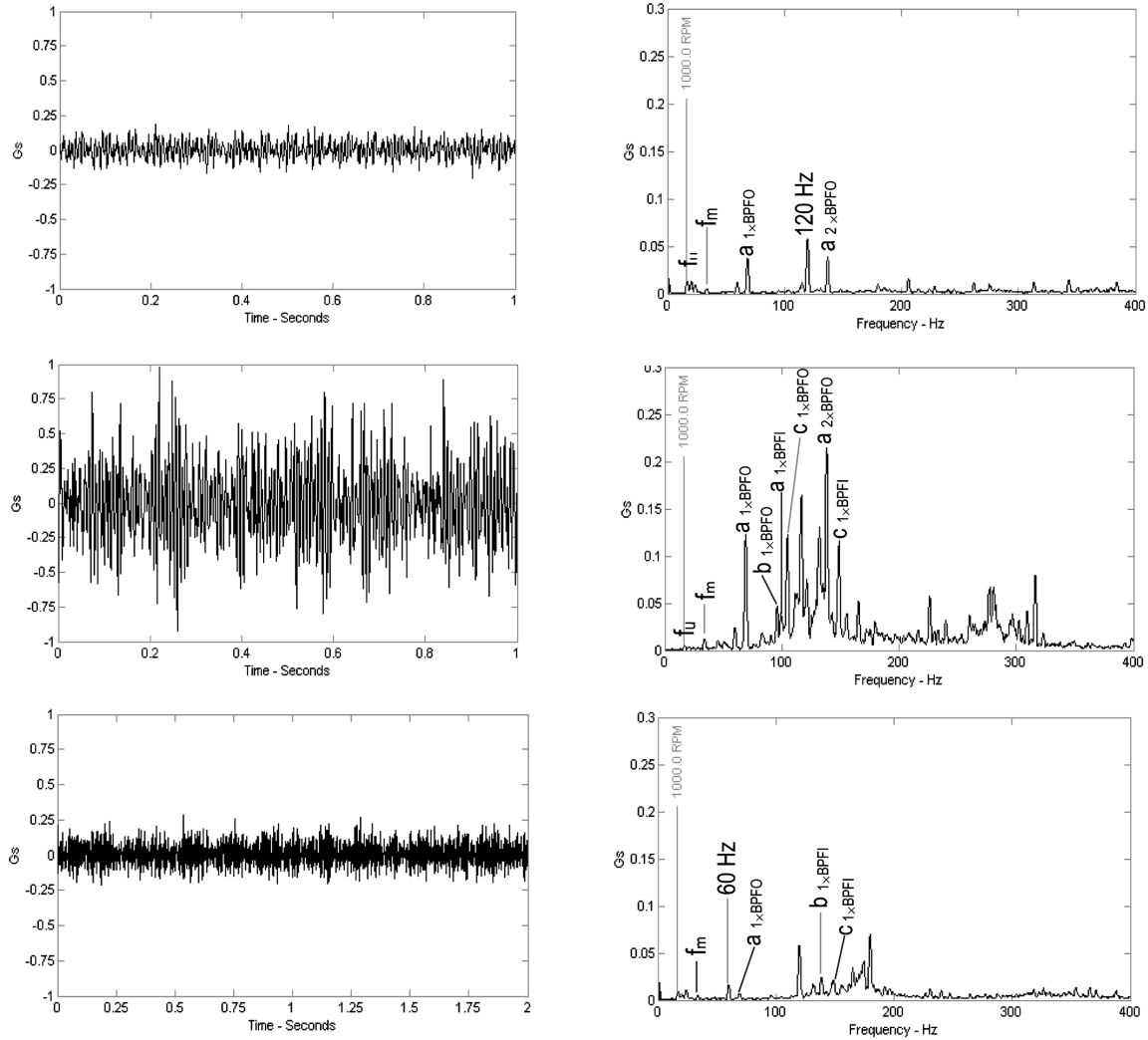
Fig. 12 Signals measured at sensor locations: S3 (top), S5 (middle) and S6 (bottom)

confirmed this finding. No harmonics of the BPFI were found in the spectrum of S2, indicating that the defect on the inner raceway must be at the incipient stage such that its signals were difficult to be detected by a dynamic force sensor (e.g. S1). On the other hand, the embedded force sensing configuration at S1 is better suited for measuring mechanical load applied to the bearing and its fluctuations, which cannot be measured directly by a shock wavebased sensor.

The sensor located at S4 is identical in design to that at S2. These two sensors were glued to the front and rear sides of the bearing housing, respectively, as illustrated in Fig. 9. The spectrum in Fig. 10 (bottom) shows a dominant effect of the bearing 6220. Since S2 is located closer to the bearing than S4, the shock wave-induced strain experienced by S2 is stronger than by S4, resulting in larger signal output at S2. Compared to S2, signals from S4 contain stronger noise. This explains that the farther away a sensor is located from the source of signal generation, the poorer its SNR, and hence the more vulnerable it will be to signal contamination.

Fig. 12 presents the signals measured by three accelerometers (100 mV/G sensitivity), located at S3, S5, and S6, respectively. It is seen that the sensor at S5, which is on the back of the bearing housing and side-by-side with S4 (Fig. 9), is most sensitive to vibrations from all the three bearings. However, it is not effective in detecting bearing unbalance (f_u). In comparison, sensor S3, located on top of the bearing housing, detected vibrations from the test bearing 6220 only. However, the S3 signal amplitude is much weaker than that from S5. On the other hand, the S3 location has shown to be more sensitive to unbalance vibration from the test bearing 6220 than S5. The overall quality of the signals detected by the sensor at S6, which is outside of the test bed cover, was very poor, and no signal component from the test bearing 6220 is seen in the spectrum. The spectrum is mainly dominated by a 60 Hz noise and its harmonics, and contains weak spectral components from the two support bearings, as shown by the 1×BPFI from bearings 2322 and 22318.

From the experimental results and analysis presented above, two conclusions can be drawn with regards to the sensor types and locations:

a) For shock wave-based sensors, location S2 delivers better result than S4;

b) For acceleration-based sensors, location S5 is more suited than S3 and S6.

To determine if a shock wave-based or acceleration-based sensor is more appropriate for bearing condition monitoring, the acceleration spectra of the two sensors (S2 and S5) were compared with each other. The second derivative was taken of the S2 signals (strain) to enable comparison with the

Fig. 13 Acceleration spectrum comparison between S2 and S5

Table 5 Signal strength from S2 and S5 at characteristic frequencies (Unit: $10^{-3}G$ or ppm)

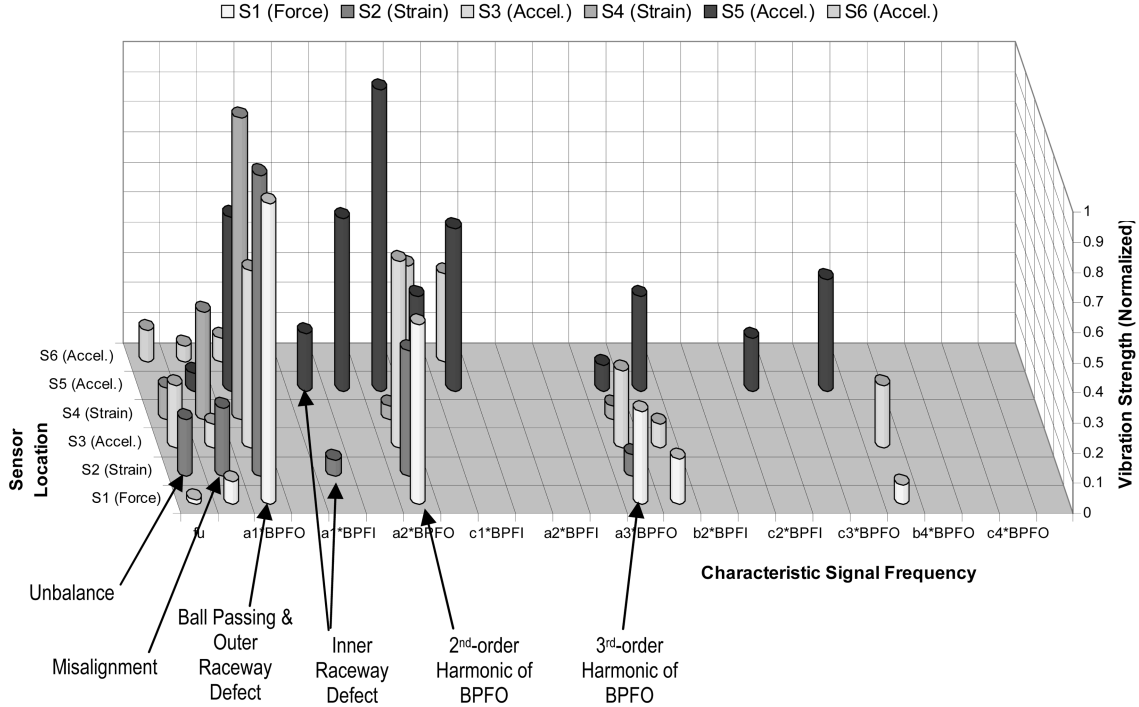| Sensor notation	Bearing	f_o	f_m	1x BPFO	2x BPFO	3x BPFO	4x BPFO	5x BPFO	1x BPFI	2x BPFI
S5 (*Accel.*)	6220		12	123	213	18	67		40	
	2232			46						67
	22318			122		79			115	37
S2 (*Accel.*)	6220		8	155	257	101	44	74	17	
	2232									
	22318									
S2 (*Strain*)	6220	1.51	1.83	8.11	3.38	0.59			0.44	
	2232									
	22318									
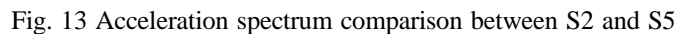
Fig. 14 Comparison of measurement results at different sensor locations

acceleration signals from S5. As shown in Fig. 13, the accelerometer at S5 was sensitive to vibrations from various sources, which not only included the test bearing 6220, but also the two support bearings and the supporting shaft. In comparison, the shock wave sensor at S2 detected mainly signals from the test bearing 6220, therefore showing a higher SNR than S5.

The signal strength at S2 and S5 is quantitatively shown in Table 5. For the S2 sensor, both the original strain and the derived acceleration measurement are listed. The unit used for acceleration is $10^{-3}G$, whereas the unit for strain is ppm.

Comparing the S2 strain signal with its second-order derivatives, it is seen that detection of the defect component at $1 \times$BPFI of bearing 6220 was not affected by the mathematical treatment. On the other hand, the derived acceleration signals have contained more "noise" components in the higher frequency band, which were not related to the test bearing 6220. This indicates that the strain signals were less sensitive to the structure-borne vibrations than its accelerometer counterpart. Another advantage of the shock-wave sensor at S2 is that it clearly indicated the unbalance and misalignment vibrations. Both of them are important parameters for bearing condition monitoring.

A summary of the signal strength detected by the sensors at locations S1 through S6 is presented in Fig. 14. The signal magnitude was normalized with respect to the maximum peak value of each corresponding sensor output. Blanks at a specific frequency indicate that either the magnitude of the peak was too small or there was no signal detected at that frequency. Depending on the sensing mechanism of the individual sensors, (dynamic force, shock waveinduced strain, or acceleration), the units used were kN, ppm, or G, respectively.

From Fig. 14, following conclusions can be drawn:

1) In general, the closer a sensor is placed relative to the bearing being monitored, the better its defect detection capability. The sensor locations S1-S5 are within the test bed cover, closer to the test bearing 6220, and thus have delivered an overall better data quality than that from S6. The location S6 is outside of the test bed cover, and the accelerometer placed at this location has shown significant signal contamination from other machine components.

2) Of the six sensors, the custom-designed, shock wave-based strain sensor at S2 has shown to be most effective in detecting vibrations from the test bearing 6220. It was able to detect vibration components related to bearing unbalance, misalignment, at f_u and f_m respectively. It also showed good signal-to-noise ratios for the BPFO's and BPFI's from the test bearing 6220, as compared to the other sensors at S1 and S3-S6.

3) The accelerometer at S5, while able to detect the defect at the inner raceway of the test bearing 6220, has shown stronger signal contamination from the two support bearings, and thus had a lower SNR for the test bearing than the custom-designed sensor at S2. 4) The custom-designed dynamic force sensor embedded into the bearing slot at S1 has shown to be effective in detecting load variations on the test bearing 6220, but was less effective than the shock wave-based sensors in detecting bearing defects, as was shown in the BPFI frequency, detected by the sensor at S2.

Research on developing a physics-based, structural dynamics-driven sensor placement strategy is being continued. Further experiments will be conducted by using bearing embedded accelerometers and shock wave sensors. Subsequent comparison from these sensors with the current sensor at S2 will reveal the best-suited sensing configuration for bearing defect detection. Furthermore, a MEMS-based sensor module is being designed for embedded bearing condition monitoring, defect diagnosis and remaining service life prognosis.

5. Conclusions

A theoretical framework for determining optimal sensor locations based on the structural dynamics of the machine system to be monitored was presented. Analysis results have shown that a sensor placed closer to the source of signal generation enables higher SNR as compared to distant sensor locations. Such a conclusion was experimentally verified by using customdesigned and commercial sensors. Key issues related to the sensor design, such as frequency response and shock wave-based sensing mechanism, were discussed. The effect of proof mass on the sensitivity of vibration sensors was systematically investigated. The effect of mass loading, and the trade-off between sensor sensitivity and effective bandwidth was also studied. It was found that high quality sensing can be achieved through a compressive shock wave sensing approach instead of the traditional acceleration measurement. The advantage of the former method is that no additional space is needed to accommodate the proof mass, resulting in a minimal space requirement for the embedded sensor, which is critical for on line, in-situ machine condition monitoring. In addition to the exemplified bearing applications, the developed integrated sensing approach can be applied to the condition monitoring of a wide range of machine systems.

Acknowledgements

References

Baydar, N. and Ball, A. (2003), "Detection of gear failures via vibration and acoustic signals using wavelet transform", *Mechanical Systems and Signal Processing*, **17**(4), 787-804.

Degtyar, A. D. and Rokhlin, S. I. (1998), "Stress effect on boundary conditions and elastic wave propagation through an interface between anisotropic media", *J. Acoustical Society of America*, **104**(4), 1992-2003.

Gao, R. and Phalakshan, P. (1995), "Design consideration for a sensor-integrated roller bearing", *ASME International Mechanical Engineering Congress and Exposition, Symposium on Rail Transportation*, RTD-**10**, 81-86.

Huang, J. S., Negnevitsky, M. and Nguyen, D. T. (2002), "A neural-fuzzy classifier for recognition of power quality disturbances", *IEEE Transactions on Power Delivery*, **17**(2), 609-616.

Holm-Hansen, B. and Gao, R. (1997), "Smart bearing utilizing embedded sensors: design considerations", *SPIE 4th Annual Symposium on Smart Structures and Materials*, Paper#3041-51, 602-610.

Holm-Hansen, B. and Gao, R. (2000), "Vibration analysis of a sensor-integrated ball bearing", *J. Vib. Acoustics, Transactions,* ASME, **122**, 384-392.

Ikolaou, N. G. N. and Ntoniadis, I. A. A. (2002), "Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets", *Mechanical Systems and Signal Processing*, **16**(4), 677-694.

Inman, D. (1996), *Engineering Vibration*, Prentice Hall, Inc., New Jersey.

Kolsky, H. (1963), *Stress Wave in Solids*, Dover Publications, Inc., New York.

Mano, R., Yoshioka, T., Miti, A. K. and Yamamoto, T. (2000), "Relationship between growth of rolling contact fatigue cracks and load distribution", *Tribology Transactions*, **43**(3), 367-376.

Mori, K., Kasashima, N., Yoshioka, T. and Ueno, Y. (1996), "Prediction of spalling on a ball bearing by applying the discrete wavelet transform to vibration signals", *Wear*, **195**(1-2), 162-168.

Nelias, D., Dumont, M. L., Champiot, F., Vincent, A., Girodin, D., Fougeres, R. and Flamand, L. (1999), "Role of inclusions, surface roughness and operating conditions on rolling contact fatigue", *J. Tribology, Transactions*, ASME, **121**(2), 240-251.

Randeraat, J. and Setterington, R. (1974), *Piezoelectric Ceramics,* 2nd edition, Mullard Company, London, England.

Rinehart, J. (1975), *Stress Transients in Solids,* University of Colorado.

Sheng, S. and Gao, R. (2003), "A wavelet-based fuzzy data fusion scheme for bearing defect severity classification", *ASME International Mechanical Engineering Congress and Exposition*, *Symposium on Intelligent Systems,* paper # IMECE2003-42584, Washington, D.C.

SKF Company (1996), *SKF Bearing Maintenance Handbook*, Denmark.

Tepfer, D. and Baldor, L. (1996), "Sikorsky says cause of crash was subcontractor's bearings", *Connecticut Post*, **5**(179), 1.

Wang, C. and Gao, R. (2003), "Wavelet transform with spectral post-processing for enhanced feature extraction", *IEEE Transactions of Instrumentation and Measurement,* **52**(4), 1296-1301.

Yan, R. and Gao, R. (2004), "Complexity as a measure for machine health evaluation", *IEEE Transactions of Instrumentation and Measurement,* **53**(4), 1327-1334.

CC