Wireless Integrated Network Sensors (WINS):

Principles and Practice

G. J. Pottie and W.J. Kaiser

Electrical Engineering Department
University of California, Los Angeles
pottie@ee.ucla.edu, kaiser@ee.ucla.edu[1]

1. INTRODUCTION

Wireless Integrated Network Sensors (WINS) provide distributed network and Internet access to sensors, controls, and processors that are deeply embedded in equipment, facilities, and the environment. The WINS network is a new monitoring and control capability for applications in transportation, manufacturing, health care, environmental monitoring, and safety and security. WINS combine microsensor technology, low power signal processing, low power computation, and low power, low cost wireless networking capability in a compact system. Recent advances in integrated circuit technology have enabled construction of far more capable sensors, radios, and processors at low cost, allowing mass production of sophisticated systems that link the physical world to networks [1-4]. Scales will range from local to global, with applications including medicine, security, factory automation, environmental monitoring, and condition-based maintenance. Compact geometry and low cost allows WINS to be embedded and distributed at a small fraction of the cost of conventional wireline sensor and actuator systems.

The opportunities for WINS depend on the development of a scalable, low cost, sensor network architecture. This requires that sensor information be conveyed to the user at low bit rate with low power transceivers. Continuous sensor signal processing must be provided to enable constant monitoring of events in an environment. By coming to decisions on these events, short message packets suffice. Future applications of distributed embedded processors and sensors will require massive numbers of devices. Conventional methods for sensor networking would present impractical demands on cable installation and network bandwidth. Through processing at source, the burden on communication system components, networks, and human resources are drastically reduced.

We limit our discussion in this paper to a security application, in which the objective is to detect and identify threats within some geographic region, and report the decisions to a remote observer via the Internet. In section 2, we describe the physical principles that lead to consideration of dense sensor networks. In section 3, we outline how energy and bandwidth constraints compel a distributed and layered signal processing architecture. In section 4, we indicate why network self-organization and reconfiguration are essential, and discuss how WINS nodes can best be embedded in the Internet. In section 5, we describe a prototype platform that enables these functions, including remote Internet control, and analysis of sensor network operation. In section 6, we present our conclusions.

[1] The authors are also founders of Sensoria Corp. of Los Angeles.

2. PHYSICAL PRINCIPLES

We now address the question of when distributed sensors are better than a single large device, given the large design cost implied in having to create a self-organizing cooperative network. We discuss in turn the fundamental limits in sensing, detection theory, communication, and signal processing that drive the design.

2.1 Propagation Laws for Sensing

All signals decay with distance as a wavefront expands. For example, in free space, electromagnetic waves decay in intensity as the square of the distance, while in other media they are subject to absorption and scattering effects that can induce even steeper declines in intensity with distance. Many media are also dispersive (e.g., via multipath or low pass filtering effects), and so a distant sensor will require costly operations such as deconvolution to partially undo the dispersion [5]. Finally, there are many obstructions that can render electromagnetic sensors useless. Regardless of the size of the array, objects behind walls or under dense foliage will not be detected.

To take a simple example, we may consider the number of pixels needed to cover some particular area at a specified resolution. The geometry of similar triangles reveals that the same number of pixels are needed whether they are concentrated in one large array, or distributed among many devices. For free space with no obstructions, we would often favor the large array, since there are no communications costs. However, in reality coverage of a large area will imply the need to track multiple targets (a very difficult problem), and further almost every security scenario of present interest involves heavily cluttered environments in which lines of sight will be obstructed. Thus, if the system is to reliably detect objects, it will need to be distributed, whatever the networking cost.

There are of course examples where is it better to concentrate the elements (e.g. radar), typically, where it is not possible to get sensors close to targets. However, there are many situations where it is possible to place sensors in proximity to targets, bringing many advantages.

2.2 Detection and Estimation Theory Fundamentals

The task of a detector is given a set of observables $\{X_j\}$ to determine which of several hypotheses $\{h_i\}$ is true. These observables may be for example the sampled output of a seismic sensor. This signal includes not only the response due to the desired target (e.g. a nearby pedestrian), but also background noise and interference from other seismic sources. A hypothesis might include the intersection of several distinct events (e.g., multiple targets of particular types present). The decision is usually based on estimates of parameters of these observations. Examples of parameters include selected Fourier, linear predictive coding, or wavelet transform coefficients, or average separation among zero crossings in the waveform. The number of parameters is typically a small fraction of the size of the observable set, and thus they constitute a reduced representation of the observations, for purposes of distinguishing among hypotheses. The set of parameters is collectively known as the feature set $\{f_k\}$. The reliability of this parameter estimation depends on both the number of independent observations and the signal to noise ratio (SNR). For example, according to the Cramer-Rao bound [6] the variance of a parameter estimate for a signal perturbed by white noise declines linearly with both the number of observations and the

SNR. Consequently, to have a good estimate of any particular feature we need either a long set of independent observations or high SNR.

The formal means for choosing between hypotheses is to construct a decision space, whose coordinates are the values of the features, and divide this space into regions according to the rule that we decide on the hypothesis h_i if the conditional probability $p(h_i|\{f_k\}) > p(h_j|\{f_k\})$ for all j not equal to i. Note that the features include environmental variations or other factors that we may measure or have prior knowledge of. The complexity of the decision increases with the dimension of the feature space, and our uncertainty in that decision also generally increases with the number of hypotheses. Thus, to reliably distinguish among many possible hypotheses, we need a larger feature space. To build the minimum size of space, we must determine the marginal improvement in the decision error rate that results from adding an additional feature. This may be as simple as including an additional term in an orthonormal expansion (e.g., FFT or wavelet analysis), or may require an entirely different transformation of the set $\{X\}$. Unfortunately, we seldom know the prior probabilities of the different hypotheses, training is often inadequate to determine the conditional probabilities, and the marginal improvement in reliability rapidly declines as more features are extracted from any given set of observables. On these facts hang many practical algorithms. We could for example apply the deconvolution and target separation machinery described in 2.1, which exploits a distributed array. Though intensive in communications and computations, it vastly reduces the size of the feature space and number of hypotheses that must be considered, as each feature extractor deals with only one target with propagation dispersal effects removed.

We may alternatively deploy a dense sensor network. Because of the decay of signals with distance, shorter range phenomena can be used (e.g., magnetics), which limit the number of targets in view at any given time (and hence hypotheses). At short range, the probability is enhanced that the environment is essentially homogeneous within the detection range, reducing the number of environmental features, and thus the size of the decision space. Finally, since at short range higher SNR is obtained, and we can use a wide variety of sensing modes that might be unavailable at distance, we are better able to choose a small feature set that distinguishes targets. With only one mode, we would need to go deep into its feature set, getting lower marginal returns for each feature. Thus, having targets nearby offers many possibilities for reducing the size of the decision space.

2.3 Communications Constraints

Spatial separation is also an important factor in the construction of communication networks. For low-lying antennas, intensity drops as the fourth power of distance due to partial cancellation by a ground-reflected ray [7,8]. Surface roughness, the presence of reflecting and obstructing objects, and antenna elevation all have an impact on propagation. The losses make long-range communication a power-hungry exercise; the combination of Maxwell's Laws and Shannon's capacity theorem together dictate that there is a limit on how many bits can be reliably conveyed given power and bandwidth restrictions. On the other hand, the strong decay of intensity with distance provides spatial isolation, allowing re-use of frequencies throughout a network.

Multipath propagation (due to reflections off multiple objects) is a very serious problem. A digital modulation will require a 40 dB increase in SNR to maintain an error probability of 10^{-5} in Rayleigh fading compared to a channel with the same average path loss perturbed only by Gaussian noise. It is possible to recover most of this loss by means of diversity. Diversity can be obtained in any of the three domains of space, frequency or time, since with sufficient separation the fade levels are independent. By spreading the information, the multiple versions will experience different fading, so that the result is more akin to the average, whereas if nothing is done it is the worst case conditions that dominate error probabilities. For static sensor nodes, time diversity is not an option with respect to path losses, although it may be a factor for jamming or other types of interference. Likewise spatial diversity is difficult to obtain, since multiple antennas are unlikely to be mounted on small platforms. Thus, diversity is most likely to be achieved in the frequency domain, for example by employing some combination of frequency hopped spread spectrum, interleaving, and channel coding. Measures which are effective against deliberate jamming are generally also effective against multipath fading and multi-user interference. This reflects the common problem of intermittent events of poor SNR.

Shadowing (wavefront obstruction and confinement) and path loss can be dealt with by employing a multi-hop network. If nodes are randomly placed in an environment, some links to near neighbors will be obstructed while others will present a clear line of sight. The higher the density, the closer the nodes and the greater the likelihood of having a link with sufficiently small distance and shadowing losses. The signals then effectively hop around obstacles. Exploitation of these forms of diversity can lead to orders of magnitude reduction in the required energy to transmit data from one location in network to another; it becomes limited chiefly by the reception and retransmission energy costs of the radio transceivers for dense peer to peer networks.

In radio systems there is thus a close connection between the networking strategy and the physical layer. The connection is even stronger when we consider the multiple access nature of the channel, since interference among users is often the limiting impairment. Management of multiple access interference is explored for example in [9].

2.4 Energy Consumption in Integrated Circuits

We now present limits to the energy efficiency of CMOS communications and signal processing circuits. Overall system cost cannot be low if the energy system must be large. A CMOS transistor pair draws power each time it is flipped. The power used is roughly proportional to the product of the switching frequency, the area of the transistor (related to device capacitance), and the square of the voltage swing. Thus, power consumption for any given operation drops roughly as the fourth power of feature size. The components that switch at high frequency and with large voltage swings dominate the power cost of the chip.

While Moore's Law implies that transistor areas will continue to decrease and thus signal processing power cost should decline with time, for radios and any communication means there are limits on the power required to reliably transmit a given distance. The power amplifier stage cannot get smaller, due to limits on the current density for semiconductors. It typically burns at least four times the radiated energy, and so will in time dominate the energy cost of radios. However, if we consider short-range communications with peak radiated powers of less than 1 mW, we continue to find that the oscillators and mixers used for up and down conversion dominate the energy budget, so that radios consume essentially the same power whether transmitting or receiving. While there will be improvement in efficiency over time, these facts suggest that networks should be designed so that the radio is off as much of the time as possible, and otherwise transmits at the minimum required level.

Processing gets cheaper with time, but is not yet free. ASICs can clock at much lower speeds and use lower numerical precision, and so consume several orders of magnitude less energy than DSPs. While the line between dedicated processors and general purpose (and thus more easily programmed) machines is constantly shifting over time, generally speaking, for computational systems which deal with connections to the physical world there needs to be a mixed architecture. The ratio in die area among the two approaches scales with technological change, and so ASICs can maintain a cost advantage over many chip generations. Convenient programmability across several orders of magnitude of energy consumption and data processing requirements is a worthy research goal for pervasive computing. In the meantime, multiprocessor systems are needed in WINS.

3. SIGNAL PROCESSING ARCHITECTURE

Security applications require constant vigilance by at least a subset of the sensors. While we desire a low false alarm rate and a high detection probability, so long as data is queued we can most of the time run energy-efficient procedures that provide high detection probabilities and poor false alarm rates. Energy thresholding and limited frequency analysis on low-sampling rate magnetic, acoustic, infrared, or seismic sensors are excellent candidates for low-power ASICs. Subsequently, higher energy processing and sensing can be invoked, if certainty levels are not high enough. Next, it might seek information from nearby sensors, for data fusion or coherent beamforming. This is a later step, since communication of raw data is very costly in terms of energy, and the procesing is also power hungry. Finally, a classification decision might be made using a large neural network or some other sophisticated procedure to provide the required degree of certainty. In the worst case, raw data may be hopped back to a remote site where a human being performs the pattern recognition. We can stop at any point in this chain where certainty thresholds are met.

Two points emerge. First, we should play the probability game to process only to the extent that we need to. Most of the time, there are no targets, and thus no need to apply our most expensive algorithm (data to humans), but there are too many circumstances in which the least expensive algorithm will fail. A processing hierarchy can lead to huge cost reductions while assuring the required level of reliability. Second, the processing hierarchy is intertwined with networking and data storage issues. How long and where data is queued depends on where in the processing hierarchy we reside; whether we communicate and to what set of neighboring nodes depends on the signal processing task. The communications costs in turn affect our processing strategy (e.g., willingness to communicate, whether processing is centralized or distributed). All of this rests on the physical constraints, and therefore the physical layer intrudes up through to applications. To illustrate, assuming a 1 GHz carrier frequency, antenna elevation of 1/2 wavelength, BPSK transmission, 10^{-6} error probability, fourth power distance loss, Rayleigh fading, and an ideal (no noise) receiver, the energy cost for transmitting 1 kilobit a distance of 100 m is approximately 3 J. By contrast, a general purpose processor having 100 MIPS/W power efficiency could execute $3x10^{8}$ instructions for this energy. Clearly, application and infrastructure permitting, it pays to process the data locally to reduce the volume of traffic and to make use of multihop routing and advanced communications techniques such as coding to reduce the energy costs.

Indeed, it is through exploitation of the application that low-power design becomes possible. For example, consider the situation depicted in Figure 1 below, for a remote security operation. These screen images display remote Internet operation of WINS. The browser screen images in (a) display events captured by an intelligent WINS NG node. For this system, the WINS node carried two sensors with seismic and imaging capability. The basic idea is that the seismic sensor can be constantly vigilant, as it requires little power. Simple energy detection can be used to trigger the operation of the camera. The image and the seismic record surrounding the event can then be communicated to a remote observer. In this way, the remote node needs to perform simple processing at low power, and the radio does not need to support continuous transmission of images. The networking allows the human (or computer) observers to be remote from the scene, and archival records to be stored. The image data allows verification of events, and is usually required in security applications that demand a human response. Both the seismic record and an image of the vehicle creating the record are shown. The WINS node and WINS Gateway node control Web pages (b-c), permitting direct and remote control (via the WINS network and Internet) of event recognition algorithms. For example, the seismic energy threshold for triggering an image can be controlled remotely. The number of images transmitted could be further reduced with an increased sensor suite of short-range detectors (e.g., infrared, magnetic), or by adding more sophisticated processing within the nodes.

Collaborative processing can extend the effective range of sensors and enable new functions. For example, consider the problem of target location. With a dense array, one means for tracking target position is for all nodes that detect a disturbance to make a report. The centroid of all nodes reporting the target is one possible estimate of the position of the target. This requires the exchange of very few bits of information per node. Much more precise position estimates can be achieved by beamforming, which requires that time-stamped raw data be exchanged among the nodes. The processing is also much more costly. However, it yields as benefits higher SNR processed data for subsequent classification decisions, long-range position location, and even some self-location and calibration possibilities for the nodes [10]. Depending on the application, it might be better to have sparse clusters of beamforming-capable nodes rather than a dense deployment of less-intelligent nodes, or it may be better to enable both sets of functions. For example, we may overlay a less dense array of intelligent nodes that can command capture of coherent data for purposes of beamforming. Allowing for heterogeneity from the outset greatly expands the processing horizons.

4. WINS NETWORK ARCHITECTURE

In contrast to conventional wireless networks, the WINS network must support large numbers of sensors in a local area with short range and low average bit rate communication (less than 1 - 100 kbps). The network design must consider the requirement to service dense sensor distributions with an emphasis on recovering environment information. We exploit the small separation between WINS nodes to provide multihop communication, with the power advantages outlined earlier. Since for short hops the transceiver power consumption for reception is nearly equal to that of transmission, the protocol should be designed so that radios are off as much of the time as possible. That is, the MAC should include some variant of TDMA. This requires that the radios periodically exchange short messages to maintain local synchronism. It is not necessary for all nodes to have the same global clock, but the local variations from link to link should be small to minimize the guard times between slots, and enable cooperative signal processing functions such as fusion and beamforming. The messages can combine health-keeping information, maintenance of synchronization, and reservation requests for bandwidth for longer packets. The abundant bandwidth that results from the spatial re-use of frequencies and local processing ensures that relatively few conflicts will result in these requests, and so simple mechanisms can be used. A low-power protocol suite that embodies these principles has been developed, including boot-up, MAC, energy-aware routing, and interaction with mobile units [11]. It indicates the feasibility of achieving distributed low-power operation in a flat multi-hop network.

It is clear however that for a wide range of applications, means must be found to conveniently link sensor networks to the Internet. Inevitably, some layering of the protocols (and devices) will be needed to make use of these standard interfaces. The WINS NG architecture design addresses the constraints on robust operation, dense and deep distribution, interoperability with conventional networks and databases, operating power, scalability, and cost (see Figure 2). WINS Gateways provide support for the WINS network and access between conventional network physical layers and their protocols and the WINS physical layer and its low power protocols. The reduction on link range through multihopping is exploited in WINS system design to provide advantages that may be selected from the set of: reduced operating power, improved bit rate, improved bit error rate, improved communication privacy (by reduction of transmit power), simplified protocols, and reduced cost. These benefits are not obtained simultaneously, but instead must be extracted depending on design emphasis.

A question that naturally arises is the extent to which Internet protocols such as TCP and IPv6 may be employed within sensor networks. While it is desirable not to have to develop new protocols or perform protocol conversion at gateways, there are several factors that demand custom solutions. First, IPv6 is not truly self-assembling; while addresses can be obtained from a server, it presupposes attachment at lower levels has already been achieved. Second, Internet protocols take little account of physical channel unreliability or the need to conserve energy; rather they focus on supporting a wide range of traffic. Embedded systems can achieve far higher efficiencies by exploiting the limited nature of the traffic.

A further question is where processing and storage should take place. As indicated in Section 3, communication costs a great deal compared to processing, and thus energy constraints dictate doing as much processing at source as possible. Further, reducing the amount of data to transmit significantly simplifies the network design, and permits scalability to thousands of nodes per Internet gateway.

5. WINS NODE ARCHITECTURE

WINS development was initiated in 1993, at UCLA. The first generation of field-ready WINS devices and software, were first fielded in 1996. (Figure 2a shows this first WINS node). LWIM-II demonstrated the feasibility of multihop, self-assembled, wireless networks. This first network also demonstrated the feasibility of algorithms for operation of wireless sensor nodes and networks at micropower level. In a joint development program with the Rockwell Science Center, a modular development platform was devised to enable evaluation of more sophisticated networking and signal processing algorithms, and to deal with many types of sensors, albeit with less emphasis on power conservation [12]. In these experiments we have come to realize the importance of separating the real-time functions that must be optimized for low-power from higher level functions that require extensive software development, but that are invoked with low duty cycle.

The WINS NG node architecture (Figure 3) was subsequently developed by Sensoria Corp. It enables continuous sensing, signal processing for event detection and local control of actuators, event identification, and communication at low power. Since the event detection process must occur continuously, the sensor, data converter, data buffer, and signal processing must all operate at micropower levels, using a real-time system. In the circumstance that an event is detected, a process may be alerted for identification of the event. Protocols for node operation then determine whether energy should be expended for further processing and whether a remote user or neighboring WINS node should be alerted. The WINS node then communicates an attribute of the identified event, for example, the address of the event in an event look-up-table stored in all network nodes. These infrequent events can be managed by the higher level processor, in this first generation a Windows CE device, chosen for the availability of low-cost tools for developers. By providing APIs that enable viewing and control over the lower level functions, the developer can be shielded from the real-time functions or delve into them as desired to make an application more efficient. Future generations will also enable plug-in Linux devices, and development is being pursued of very small but limited sensing devices that can interact with WINS NG nodes in heterogeneous networks, for example as intelligent tags. Such small devices might scavenge their energy from the environment by means of photocells or piezoelectric materials to capture energy from vibrations, to achieve indefinite lifetimes. A clear technical path of increased circuit integration and improved packaging exists to make very low cost and compact devices in the near future.

6. CONCLUSION

We have presented an overview of the physical considerations that lead to the design of densely distributed sensor networks, and reviewed the advantages of layered and heterogeneous processing/networking architectures for these applications. The close intertwining of processing of networking is a central feature of systems that connect the physical and virtual worlds. Development platforms are now available that will more easily enable a broader community to engage in fundamental research in networking and new applications, advancing us towards truly pervasive computing.

7. REFERENCES

1. K. Bult, A. Burstein, D. Chang, M. Dong, M. Fielding, E. Kruglick, J. Ho, F. Lin, T.-H. Lin, W. J. Kaiser, H. Marcy, R. Mukai, P. Nelson, F. Newberg, K. S. J. Pister, G. Pottie, H. Sanchez, O. M. Stafsudd, K. B. Tan, C. M. Ward, S. Xue, J. Yao, "Low Power Systems for Wireless Microsensors", *Proceedings of International Symposium on Low Power Electronics and Design,* pp. 17-21, 1996.
2. M. J. Dong, G. Yung, and W. J. Kaiser, " Low Power Signal Processing Architectures for Network Microsensors" *Proceedings of 1997 International Symposium on Low Power Electronics and Design*, pp. 173-177, 1997.
3. T.-H. Lin, H. Sanchez, R. Rofougaran, and W. J. Kaiser, "CMOS Front End Components for Micropower RF Wireless Systems" *Proceedings of the 1998 International Symposium on Low Power Electronics and Design,* pp. 11-15, 1998.
4. G. Asada, M. Dong, T. S. Lin, F. Newberg, G. Pottie, H. O. Marcy, and W. J. Kaiser, "Wireless Integrated Network Sensors: Low Power Systems on a Chip" , *Proceedings of the 24th IEEE European Solid-State Circuits Conference*, 1998.
5. T. Yu, D. Chen, G.J. Pottie, K. Yao, "Blind Decorrelation and Deconvolution Algorithm for Multiple-Input, Multiple-Output System," *Proc. SPIE* vol. 3807, 1999.
6. H. Van Trees. Detection, Estimation and Modulation Theory. Wiley: New York, 1968.
7. T.S. Rappaport, Wireless Communications: Principles and Practice. Prentice Hall: 1996.
8. K. Sohrabi, B. Manriquez, and G. Pottie, "Near-ground wideband channel measurements," VTC '99

9. G.J. Pottie "Wireless Multiple Access Adaptive Communication Techniques" in *Encyclopedia of Telecommunications*, vol. 18, Ed. Froelich/Kent. Dekker: 1999.
10. K. Yao, R.E. Hudson, C.W. Reed, D. Chen, F. Lorenzelli, "Blind Beamforming on a Randomly Distributed Sensor Array System," *IEEE J. Select. Areas in Comm.*, vol. 16, No. 8, Oct. 1998, pp. 1555-1567.
11. K. Sohrabi, J. Gao, V. Ailawadhi, G. Pottie, "A Self-Organizing Sensor Network," *Proc. 37th Allerton Conf. On Comm., Control, and Computing*, Monticello IL, Sept. 1999.
12. J.R. Agre, L.P. Clare, G.J. Pottie, N.P. Romanov, "Development Platform for Self-Organizing Wireless Sensor Networks" *Aerosense'99*, Orlando FL, April 1999.

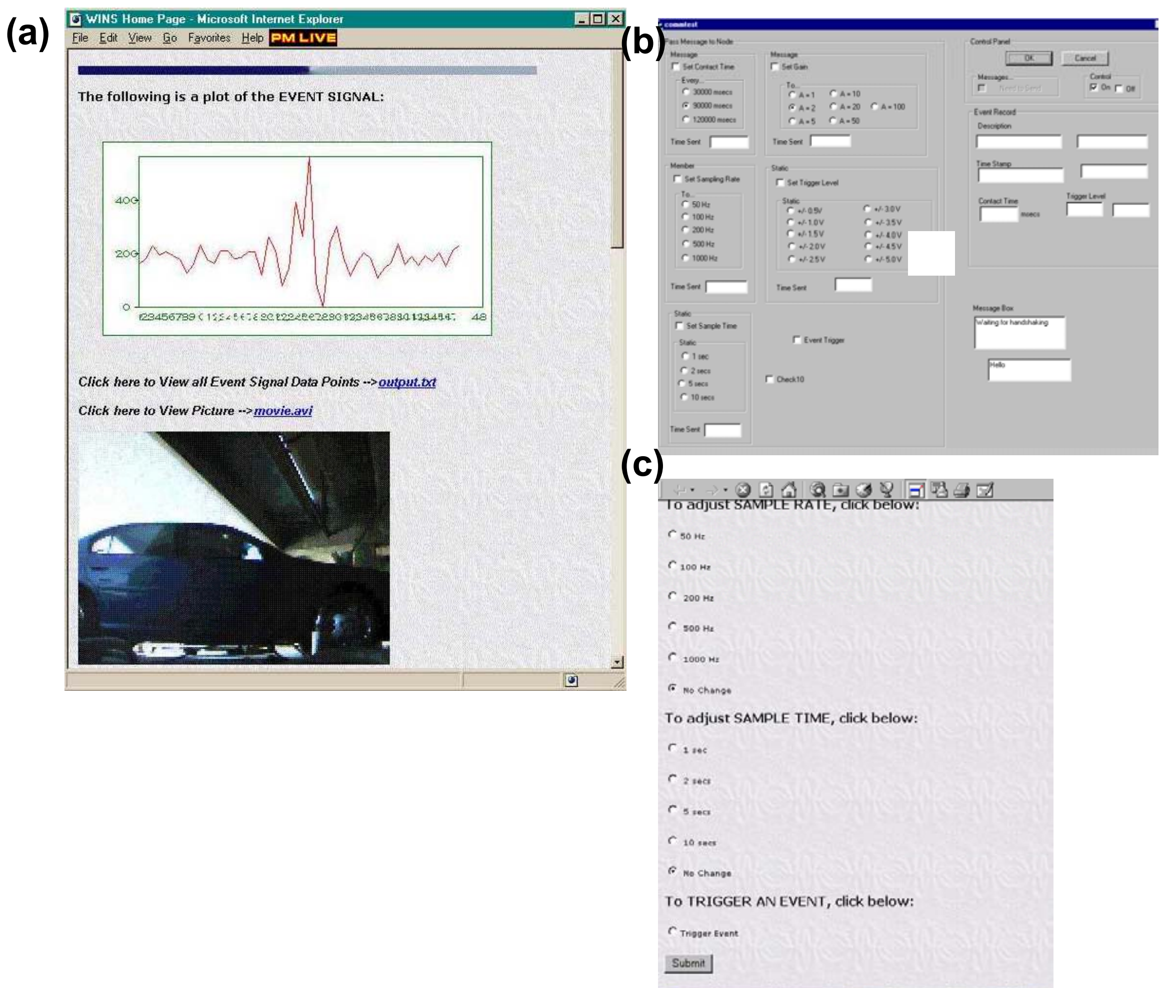

Figure 1. Remote Access to Seismic and Image Data

(a) **(b)**

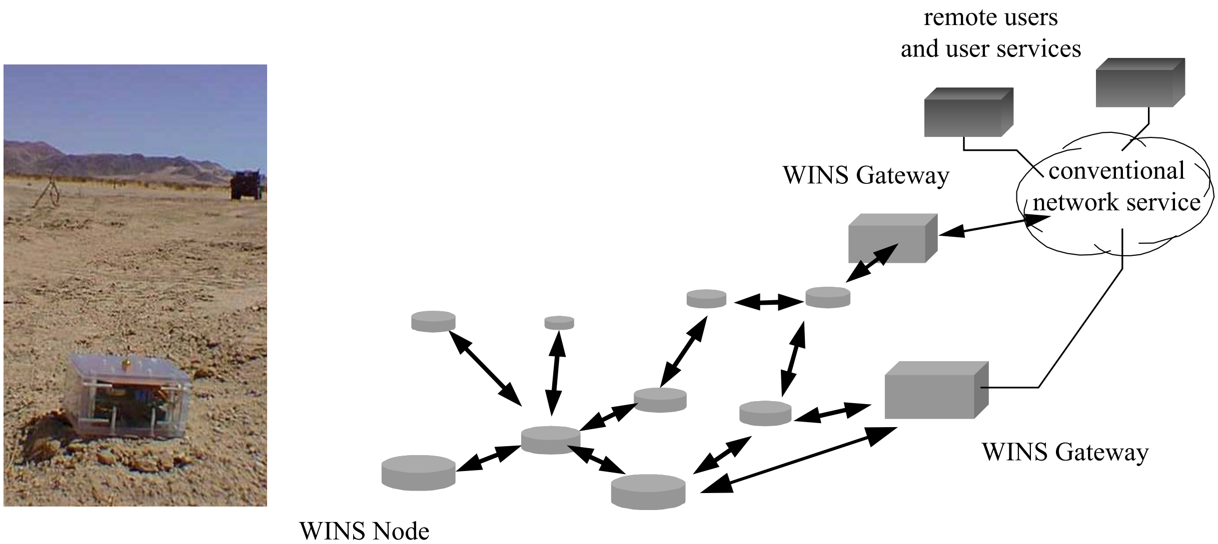

Figure 2. WINS Network Architecture.

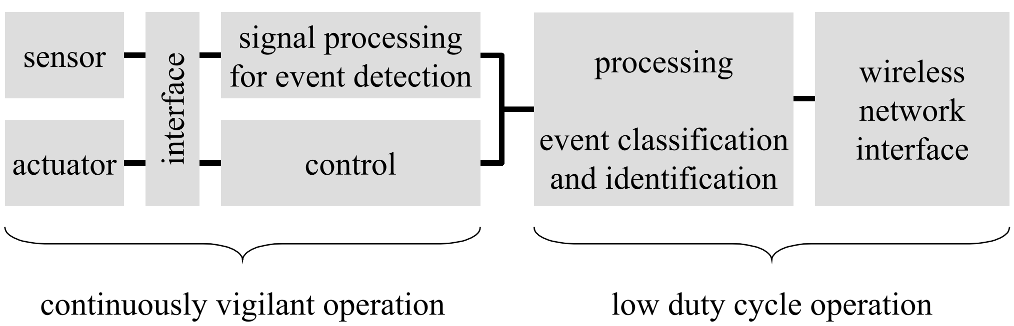

Figure 3. WINS Node Architecture.