Project Description

1. Introduction
          We propose a broad-based research effort to develop a large, ad-hoc, wireless network of millimeter size smart sensor nodes. While there will be various novel aspects of this network, an especially unique attribute will be the implementation of these smart, wireless sensors using 3-dimensional (3D) integrated semiconductor process technology. The 3D integration will facilitate the development of smart, wireless sensors that can be both very small and, at the same time, be functionally robust. In summary, the micro-sensor network will contain anywhere from ten to perhaps thousands of these 3D integrated smart millimeter-sensor nodes. Once these millimeter-nodes are distributed, they will self-assemble into a wireless network. The nodes will then sense environmental attributes, communicate these attributes to other nodes in the network, fuse their data, perform distributed computations and then reach global decisions.
          The generic operational structure of a network node consists of one or more analog sensors which take data from the environment. This information is then digitized and stored in the microprocessor. The data stored in the microprocessor is then distributed wirelessly to the network using the transmitter. The receiver part of the smart sensor node obtains information from the other nodes in the network. This information is then stored in the microprocessor as well. The microprocessor serves two major functions. First, it processes and helps fuse the data from its own sensors with the data it receives from the rest of the network to help in the distributed network decision making. In addition, the microprocessor must control the communication algorithm for the network. In other words, it must tell its own transceiver when to receive information from other nodes, and direct when to transmit its own data and decisions. The network will self-assemble. Initial prototypes will communicate digitally using a time-division multiple access (TDMA) scheme, with data transmitted using frequency shift keying (FSK). A functional diagram of a network node is given in Figure 1.
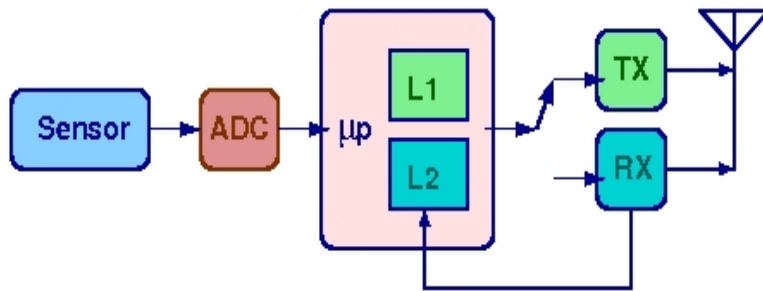

Figure 1
Smart Wireless
Sensor functional
diagram

          A number of broad research efforts in the development of sensor networks technology and systems have been undertaken in recent years. Among these are the PicoRadio project[1] the mAMPS project[2]; and the Center for Embedded Networked Sensing[3]. Other interesting approaches have utilized unique sensor structures[4].
          Our ad-hoc sensor network moves in a different direction with the following unique attributes:
- Very small geometry nodes enabled by 3D integration.
- A uniquely low power approach to communication IC's using regenerative transceivers.
- Novel large energy storage capability enabled through the use of extremely high-k dielectric materials and interdigitated 3D capacitors, on chip high Q, 3D inductors and transformers.
- Unique antenna structures that allow for transmission in the low GHz region, while at the same time being only millimeters in size.
- We are developing unique signal processing algorithms that already allow for object location detection tracking in very energy efficient ways.
- We have unique physical modeling capabilities to help predict the performance of 3D IC's, particularly with respect to thermal noise and heat dissipation, and electromagnetic coupling within the network and for 3D sensors themselves.

- Leveraging our previous work parameterized local search algorithms (PLSA) on evolutionary algorithms for multi-dimensional optimization of signal processing software, we will explore new global search strategies that work effectively for the PLSAs, design constraints, and parameter spaces associated with our proposed class of millimeter-size sensor networks.
- We have already developed a wireless smart sensor network to form a test-bed for our research.

In summary, we plan to design and prototype an advanced wireless network. The network will be composed of smart sensor nodes which are fabricated using 3D IC technology. Using 3D integration capabilities, as well as other novel contributions including ones listed above, we will design and fabricate a network that has considerably more functionality than is possible with standard planar IC technology.

1. 2. Proposal Structure

First, we describe the chronological development by dividing the program into three phases. We then describe our background work, where we have already developed a working test-bed wireless sensor network. Next, we discuss the various hardware components of the sensor nodes that we are developing, including transceiver design, digital design, novel materials, antennas, and power sources, and their implementation using 3D integration. Finally we discuss signal processing algorithms for data collection and fusion, and distributed decision making, as well as energy efficient communication algorithms tailored for smart sensor networks.

1. 3. Planned Program Chronology

Our proposed research will be divided into three major phases as described below:

Phase 1: In the first phase, we will prototype working networks using commercial off the shelf components. This network will serve as a testbet for our network communication and sensing algorithms. In fact, as we will discuss in our background research below, we have already developed such networks. Here we will improve on our current designs, and determine how to optimize networks using current off-the shelf integrated circuits. Sensor nodes will be prototyped on printed circuit boards we design with Eagle software, and built using PCB fabrication services. The nodes will be programmed with application specific algorithms, and the sensor network will be established. The first network will be programmed to first sense a global maximum, eg, temperature. The next distributed network task will be to detect motion across a boundary. After initial deployment of the network, we will continue to reprogram the network with new algorithms for different applications. We will also develop algorithms to improve the resolution in time and space of the network. We will continue to utilize different sensors and work them into the network, including temperature, vibration, acoustical, optical and chemical sensor applications.

Phase 2: We will develop our own sensor nodes using a combination of existing IC's, and ones that we build and design ourselves. Our own designs will be developed using CMOS technology, designed with Cadence, and fabricated through MOSIS. The nodes will be assembled into planar, multi-chip modules. The multi-chip modules will be prototyped using a 100-micron resolution pick and place machine. The nodes will be programmed and the sensor network will be established.

Phase 3: We will transform the planar multi-chip module smart sensor nodes into our millimeter sized wireless smart sensor nodes using 3D integration of integrated circuits. The 3D integration will be achieved by first designing and having our own chips fabricated through MOSIS. The chips will be specially designed for 3D implementation, requiring development of new 3D design rules and establishing of specific symmetries for 3D interconnects between chips. 3D interconnects will be facilitated by first thinning the MOSIS chips to 20micron thickness, then applying the 3D vertical interconnect process that we developed previously in collaboration with the Laboratory of Physical Sciences.

Throughout each phase we will be developing new algorithms to increase the sensitivity, resolution and robustness of the network. We will also develop new unique physical designs and hardware to improve network sensitivity, power efficiency, longevity and connectivity. To achieve this, we divided our research team into two basic groups: the applied math group and the applied physics group. The applied math group will be responsible for algorithm development, while the applied physics group will focus on novel and robust development of network smart sensor nodes.

2. Wireless Sensor Network Test-bed

As part of the background for the proposed work, we have already developed a small-size ad-hoc sensor network. The development of this network was supported elsewhere, and its main function is to serve as a test-bed for implementing and comparing different communication and data processing algorithms, for a variety of applications. It will also serve as a test-bed for optimizing the locally constructed multiple access and data fusion algorithms we plan to develop, and for evaluating their performance characteristics in performing sample tasks over large-scale ad-hoc sensor networks. The network test-bed will also serve as a vehicle for comparing various protocols, including the Zigbee scheme. We are fortunate to have this project underway, as it can readily complement our proposed NSF work: *Autonomous Sensor Networks through 3D Integration, Power-Efficient Algorithms, Optimized Configurations and Advanced Materials.* Using the test-bed will enable us to determine how 3D integration and advanced physical design will allow us to improve or best utilize the aforementioned existing network sensing and communication.

We have currently developed a TDMA based wireless network using the peer-to-peer topology for transmission. Each network node comprises a PIC 16F88 microcontroller (MCU) made by Microchip, and a TR-916-SC-P transceiver manufactured by Linx Technologies. The network self-assembles by first having nodes transmit their respective clock references. Through locally negotiated clock-reference readjustments, the clocks of the individual nodes clocks achieve synchronization, and establish a global clock reference. Subsequently, nodes continuously read analog sensor input, convert the information to digital format. The information is then packaged for digital transmission using start and stop bits, and error correction by the microprocessor. These data packets are transferred to the transceiver and transmitted to the network at the proper time slot. When in receive mode, nodes accept data from the transmitting node and store the information for that specific node. The network is currently programmed to detect the maximum of the data broadcast by all the sensors in the network. For example, if the network were sensing temperature, it would make a distributed decision as to which node was the hottest. While sensing which node is at the maximum temperature can be relatively straightforward, we are now extending the network to identify objects, sense their position, as well as their motion. This will be applied to numerous other problems to determine its applicability and limitations. In doing so, we will help ascertain the strengths of TDMA based systems.

We plan to continually revise and upgrade our test-bed network in various ways. First, we will continue to experiment with different algorithms to establish the TDMA protocols. These experimental algorithms, which are discussed later in the proposal, will usually be implemented in software by reprogramming the node's micro-controller. In addition to comparing various algorithms that are implemented in software, we will use the test-bed to implement different hardware schemes, especially the Microchip 18 series MCU and Chipcon CC1000 and CC2000 series transceivers and transceiver/MCU.

3. 3D Integration: Process Flow

We have developed a 3D integration process technology flow in collaboration with the Laboratory of Physical Sciences in College Park Maryland[5,6]. The most challenging fabrication process in current stacked circuits technology is through-hole 3DI. The fabrication relies on advanced back-end semiconductor processing and packaging technologies. Back-end processing is used primarily for the creation of high aspect ratio (HAR) through-wafer/die vias. Packaging work includes die de-packaging, substrate thinning, and die placement.

Commercial or custom die are obtained. Chips are then cleaned using standard IC processing. Once prepared, the die are then accurately picked and placed onto a carrier wafer using an automated flip-chip tool that dispenses a temporary adhesive such as a UV-curable epoxy. This procedure ensures that the die are aligned in X and Y directions within 5 microns. In addition, it is critical that the die have a uniform Z-height across the carrier wafer. After the placement process, the die are thinned using a combination of standard wafer grinding, single-sided chemical etching, and Chemical Mechanical Polishing (CMP) processes. These steps result in die with a thickness of less than 20 microns and a total thickness variation (TTV) of less than 1 micron.

Immediately following the thinning process, the thinned die are cleaned using standard IC processing and exposed to a gas plasma to prepare the surfaces for covalent bonding. Covalent bonding relies on water vapor to create temporary bonding via van der Walls forces by initiating a H_2O chain between the die and the handle wafer. When the bonded materials are cured at an elevated temperature (above 150˚C) the hydrogen atoms diffuse through the silicon and an atomically smooth, very thin SiO_2 interface is grown to provide a permanent bond after which the adhesive material is removed to free the re-integrated die wafer to which subsequent layers are bonded.

When a 2-layer die stack is formed, HAR through die vias are dry etched, and metallized to create an electrical connection between the two die surfaces, as illustrated by the cross-section in Figure 2. The bonding and interconnect processes are repeated for each layer to form the final stacked circuit.
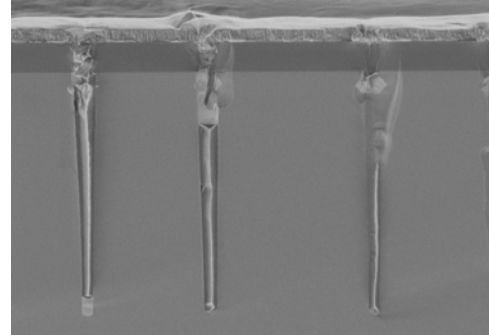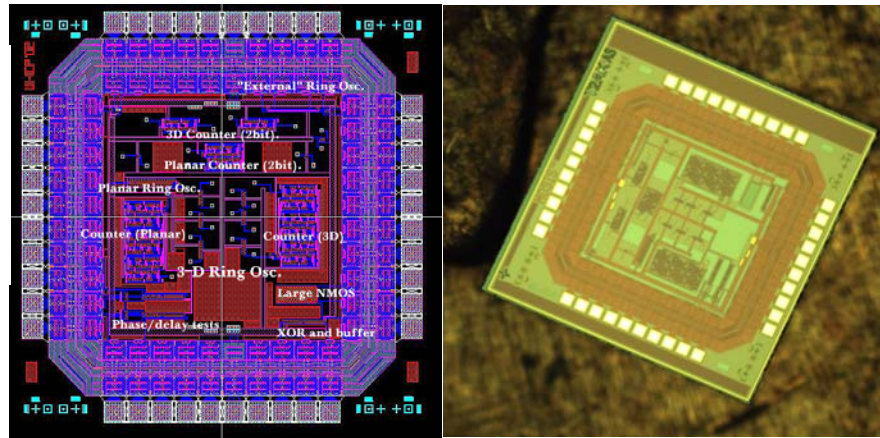
We have also had a clock circuit intended for 3D integration fabricated, as displayed in Figure 3.

Figure 2 20 SEM cross-section of HAR vias

Figure 3 Layout and final die photograph for a custom-designed chip for 3D integration.

In addition to the process flow, we have also developed unique computer modeling capabilities for analyzing parasitic effects in 3D IC's related to electromagnetic coupling and thermal effects[7-10].

4. Radio Frequency VLSI Hardware

To prototype 3D ICs sensor nodes, we eventually have to develop our own integrated circuits that are compatible with 3D integration, and are optimized for smart sensor networks. We plan to develop two types of transceiver circuits. One will be based on the more conventional modern PLL based FSK transceiver. The other will be based on the super-regenerative transceiver concept, using ASK. This is a more experimental design that we plan to implement to significantly reduce power consumption. The design requires very few components, and thus appears to be ideally suited for extremely low power autonomous networks.

The smart sensor transceiver system will have to be low power, low noise, and have physical dimensions in the millimeter range. In addition, the transceiver must be relatively inexpensive to fabricate so that thousands can be deployed without prohibitive costs. The frequency of operation must be in the gigahertz range, so that commensurately sized antennas may be utilized. The design will need to accommodate the TDMA multiplexing scheme using either FSK or ASK modulation in order to establish the ad-hoc network described in Section 9 of this proposal. Also, the design should be well suited for 3D integration. Incorporation of 3D integration transcends the current state of the art, which mainly relies on 2D integration: this should facilitate the development of unique compact smart sensor systems. The size, frequency, modulation, power, 3D integration and cost criteria all argue for the use of CMOS technology

to develop and fabricate smart sensor nodes. In addition, reduction in size also corresponds to commensurate reduction in supply voltages, and for various components of the system, a quadratic reduction in power.

4. 1. Super Regenerative Transmitter/Receiver

Single-Stage, Low Power, Microwave Transceivers: Perhaps the simplest radio-frequency transmitter/receivers are based on single stage amplifiers operating with some positive, or regenerative, feedback[11]. These topologies are as old as the consumer radios of the early twentieth century, which use an amplitude modulation format. They were developed when it was noticed that a coupling between the gate and anode of a tuned amplifier using a vacuum tube could cause high-gain and good selectivity.
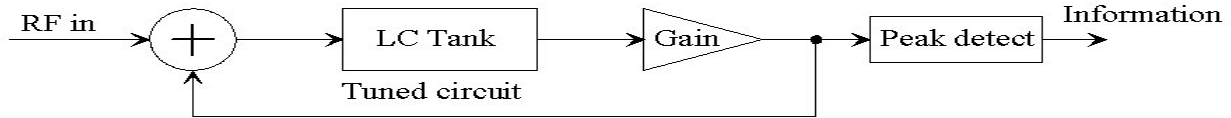

Figure 4 Block diagram of the super-regenerative receiver.

The output of the tuned circuit is amplified and fed back to mix with an RF input signal, as shown in Fig. 4. The positive feedback causes the output amplitude of the gain block (the amplifier) to vary with the RF input. This variation is fed to a peak detector circuit to retrieve the transmitted information.

Figure 5 shows how the same basic topology can be used to transmit a signal. The high-frequency output of the amplifier is now used to drive an antenna while the "Information in" signal is used to modulate the amplitude of the transmitted RF. The benefits of using this architecture are simplicity, low power (extremely important in this application), and high-frequency operation. The concerns with using a regenerative topology are tuning and stability. Because the small sensor nodes are autonomous, the lack of tuning ability is a concern. We believe that the absolute value of the carrier frequency used by the sensor nodes for communication is not important but rather what is important is that they all use the same frequency for communications. This synchronization between nodes should be accomplished by using nodes fabricated in the same process run. Control of the stability will be investigated using a super-regenerative architecture [11]. In this architecture a "quench oscillator" is used to damp the oscillations propagating through the positive feedback path.
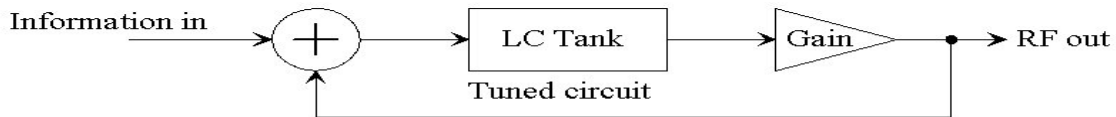

Figure 5 Modifying the block diagram in Figure 4 for transmission.

We have been investigating the use of CMOS technology [12] in RF applications. Figure 6 shows one of the many RF test structures fabricated and tested in the RF lab at Boise State University.
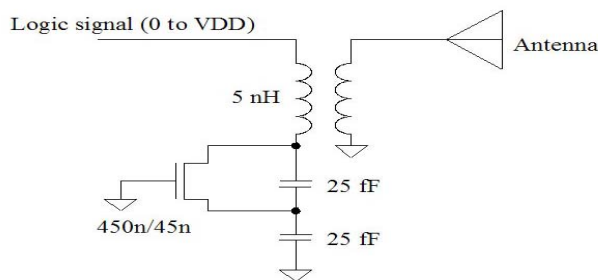

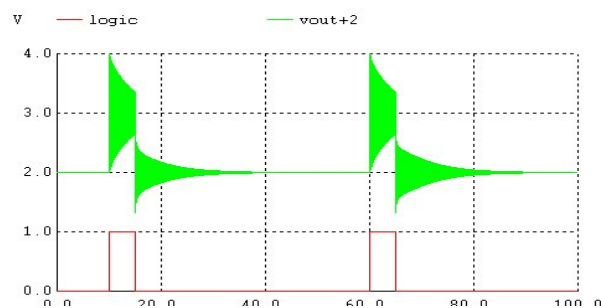*Figure 6* Transmitter topology. Components selected for a 20 GHz carrier frequency.

Figure 7 The controlling logic and transmitter output signals (20 GHz carrier).

The proposed approach, while simple, has a very high probability of success and can be used for a first generation implementation of millimeter scale smart sensors. Initial SPICE simulations of the circuit in Figure 6 show a very encouraging performance. A short 20GHz RF pulse is generated, which naturally squelches. These results are shown in Figure 7. This signal represents the carrier for our ASK transceiver. Initial designs will have digital levels detected using a simple Schottky diode/inductor structure, which we have fabricated through MOSIS[13].

4.2. FSK Transceiver Topology:

Our alternate transceiver design will be based on the modern conventional PLL approach. The transceiver will be narrow band, containing receiver and transmitter blocks. Initial designs will be centered to operate at carrier frequencies in the low GHz range. (We are investigating FCC regulations, but will probably establish the network to operate at an IMS band.) Noise levels at the input should be small enough so that receiving small sensors can detect signals which are at least as low as -100dBm. Precise values of these numbers will depend on the technology, power, noise levels, and the details of the TDMA communication network we develop. The receiver hardware will consist of a low-noise amplifier (LNA), a mixer, an IF amplifier, and a local oscillator (PLL-frequency synthesizer). The transmitter will require a frequency synthesizer, a power amplifier and will share the

Figure 8 PLL chip designed for FSK modulation

PLL with the receiver. We have already developed FSK transceivers based on the PLL topology. Figure 8 shows the chip layout. Since we will be using TDMA, we will not require especially stringent phase noise restrictions from the PLL, so basic designs will be used. A major concern will be how to minimize power, while also minimizing noise. This will be achieved using supply voltages as an optimization variable.
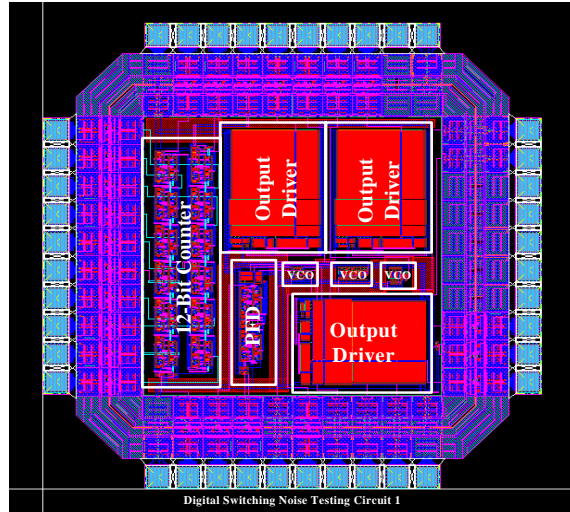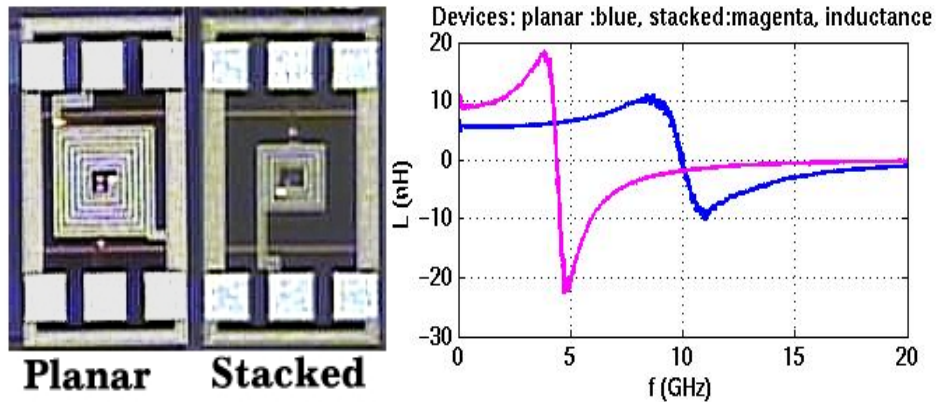
4. 3. Passive Structures (Inductors and Capacitors)

Passive components will present a challenge with respect to size constraints. Inductors will be in the nanohenry range, while capacitors will be in the tenth (0.1) picofarad range. 3D integration is great asset in fabricating these passive components because the added volume will greatly facilitate fabrication of these elements while at the same time minimizing their resistive parasitic elements which reduce the quality factor. 3D integration also provides an added degree of freedom, letting us maximize inductance while minimizing intrinsic parasitics.

Figure 9 Layouts and measured inductance for a planar and a 3D inductor.

We have investigated different integrated inductor and transformer designs [14]. We have fabricated a planar inductor and a stacked inductor using the MOSIS service. Our measurements indicate that the 3D inductor displays a higher

inductance up to its self-resonant frequency while taking up a far smaller chip area. Figure 9 displays die photos and our measurement results of the fabricated planar and stacked inductors. We are also developing algorithms that allow us to take advantage of this extra degree of design freedom. One of these algorithms uses nonlinear best fit methods to minimize parasitic capacitance, resistance, and chip area for a given required inductance. This results in a skewed vertical inductor stack.

Similar algorithms will be developed to take advantage of naturally self resonating LC structures.

We also plan to apply our background research on carbon nanotubes to examine their use as inductors. Our background investigation indicates that CNTs display negative differential mobility, which may give rise to large kinetic inductance, and require a small volume, thereby saving considerable IC real estate[15,16]. We are developing methods to incorporate them into the 3D integration process flow.

5. Physical Design: Power Sources Using 3D Integration and High K Dielectrics

Power generation is an enabling technology for distributed sensor networks. Our ability to generate power by any means is proportional to the volume of the power generator. Distributed sensor systems represent a unique challenge in that they are of small volume and they cannot be "plugged in" to external power systems. While laptop computers make use of "low power" electronic design, their batteries last for times on the order of hours and these computers are about three orders of magnitude larger in volume than the distributed nodes envisioned here. Distributed networks may be expected to last anywhere from days to essentially permanent installation.

There are two approaches to power generation for distributed networks. These are "internal" and "external" generations. Internal generators are batteries and fuel cells. External generations require external power sources 'beamed' to the requisite node. These include solar cells, RF or laser generators. It may also be possible to move a microwave generator through the field using land-based transportation in some cases.

For this program, we seek to merge these two types of technology. We propose utilizing an electrochemical cell (Ruthenium oxide/sulfuric acid), chargeable through a "rectenna" approach - microwave power beamed from a unmanned autonomous vehicle (UAV), rectified and stored on the battery plates [17].

The advantage of the electrochemical cell is that the acid space-charge – the Helmholtz double layer – effectively forms a pinhole free dielectric which enables charge storage as well as acting as a current source. This layer is thin – on the order of nanometers. It is essentially "self assembling." That is, it is not the result of any deposition or "spin on" process. As such, it is guaranteed to be continuous. The thin layer provides a high capacitance for charge storage. The recharging microwave beam stores relatively large amounts of energy on the electrodes as well as reversing the direction of the power generating chemical reaction.

Our work is unique in a number of ways. While proof-of-principles demonstrations exist in the open literature for each of the approaches listed, this will be the first time a number of the technologies described above will be fused together into a single system. In addition, we will use silicon micromachining techniques to "corrugate' the battery surface with square pillars, enhancing the surface area in an amount equal to the aspect ratio of the pillar (the height of the pillar ratio'd to the square side dimension), illustrated in Figure 10.

Metal

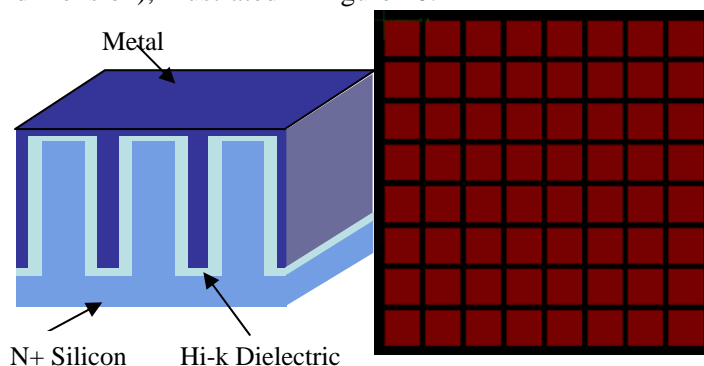N+ Silicon   Hi-k Dielectric

Figure 10
"Corrugation" in silicon produced by deep reactive ion etching, and the mask for the process

The whole corrugated structure will be rendered conducting by phosphorous diffusion, sputter-coated with ruthenium oxide and submerged in a shallow troth of sulfuric acid. The acid volume is about 1 drop! This will be sealed using a

gold-coated (acid resistant) lid. The lid becomes the battery/capacitor electrode, and the substrate contact becomes the counter-electrode. The process is extremely simple – only one photo-mask step is required. The 'self-assembly" aspect of the effective dielectric formation is aimed at providing high yield.

6. Design of Physically-Small Antennas

Antennas represent a fundamental building block in any sensors network. Unfortunately, in size, antennas lag behind other sensors-involved technologies such as the electronic components needed to process the data or the primary sensing modules. Antennas remain much larger than any other components in sensor modules operating in the lower end of the microwave spectrum. Antennas must meet specific design constraints such as polarization, power gain, efficiency, and most importantly, physical size or volume. The fundamental limitations of electrically small antennas are well documented, however, what constitutes a *small* antenna is something which is not clearly defined. Recent advances in material technology and negative index material strongly supports this assertion. A primary objective of our work in this proposal will investigate the potential of *physically- small* antennas, more specifically, we will focus on an important class of low-profile antennas of the microstrip variety. Our work will be divided into the following subgroups:

(a) Design of Printed Antennas Using High Permittivity Material

Miniaturized antennas are a major objective when designing physically small sensors that operate around the 1-5GHz region. The advantage of operating over this band is the affordability of the technology for oscillators and other components. By miniaturizing an antenna, we mean making the antenna very small in comparison to the wavelength in free space at the frequency of operation. For instance, if we consider a microstrip patch antenna operating at 1GHz, the length of the antenna would be close to half of the wavelength at this frequency, corresponding to approximately 15cm. Clearly a sensor having one of its components close to this size is impractical. If we can reduce the size of this patch antenna by one or two orders of magnitude, then the antenna size will be approaching the size of a sensor that fits in a sphere of 1cm in diameter. We propose to introduce high permittivity material, typically referred to as high-k material, into the design of physically but not necessarily electrically-small antennas [18-21].

Two approaches will be pursued here that will incorporate high-k material in the design of antennas. (Figure 11 shows preliminary data for the relative dielectric permittivity of BaTiO3 as a function of temperature). The simplest approach in doing this would be to include multi-layer substrates and/or using multi-layer superstrates. The primary motive for using high-k multi-layer substrate is reducing the actual length of the wavelength within the microstrip antenna cavity. The primary motive for the use of multi-layer superstrate, on the other hand, is to effectively impedance-match the fringing fields (of the cavity)

Figure 11 Low-frequency dielectric constant of sputter-deposited BaTiO3 as a function of temperature (film thickness was 200 nm).

which in essence give rise to the antenna radiation. In the second approach, we plan to design substrates with adjusted permittivity by inserting small cylindrical rods of high-k material into grooves in a relatively low-k material. This work will theoretically investigate the potential of these approaches and investigate the effect of material losses in the dielectric on the gain and efficiency of the antenna. Different topologies will be studied, numerically simulated, constructed at the UMCP facilities and tested.
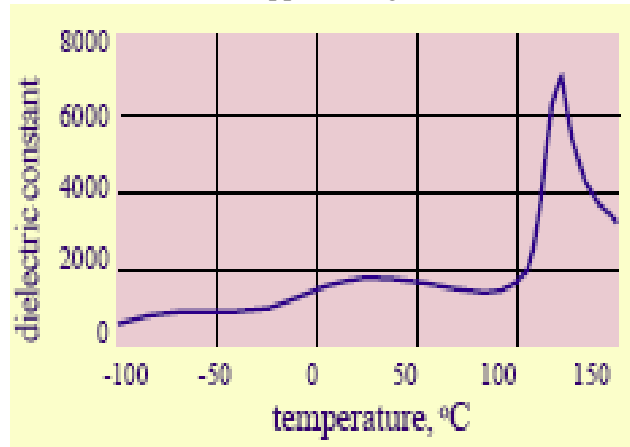
(b) Robust Optimization of Small Antennas

Determining the optimal geometric configuration for microstrip patch antennas is a challenging task since the relationship between the physical parameters of the antenna such as substrate depth, feed topology, strip length and overall performance is complex. There are various design approaches reported in the literature. An approach that was proven to be highly suitable to multi-objective optimization is the Genetic Algorithm (GA), which is part of a larger class of evolutionary optimization techniques. In previous works, it was shown that when using GA, optimal antenna performance can be achieved; however, the topology of the final design was either difficult to manufacture, or would require very specific material parameters. Such constraints can render the final optimal design impractical to implement or prone to appreciable variation in performance when some of the design parameters are slightly altered.

In this work, we proposed to investigate a novel robust optimization approach to determine the parameter combinations that results in optimal antenna performance. Robust optimization is intended as an optimization tool capable of reducing the sensitivity of an optimum design solution to uncontrollable parameters variations. As a demonstration, a microstrip antenna was designed using a commercially available electromagnetic full-wave simulator and a robust GA was concurrently applied to determine the different physical parameter combinations that result in optimal antennas performance. The robustness of this approach was then evaluated by perturbing the computed optimal parameters until the performance variation reached acceptable boundaries.

7. Integrated Sensors 3D Smart Wireless Nodes

Recent developments in MEMS and self-assembly have produced an array of new sensors suitable for 3D integration into distributed sensor nodes. One of the major outcomes of our sensor network will be a system capable of tracking *and recognizing* an on-coming threat from its acoustic signature. To this end, miniature microphones will be required. Also, the threat of toxic gas and harmful biological agents is of paramount concern to homeland security. Chemically modified self-assembled monomers can be integrated with integrated electronics to create compact sensors of high sensitivity. Both of these sensor types will be studied in the proposed work.

MEMS microphones are both sensitive and compact. They are extremely simple structures, but very hard to build! A typical MEMS acoustic sensor is shown in Figure 12.
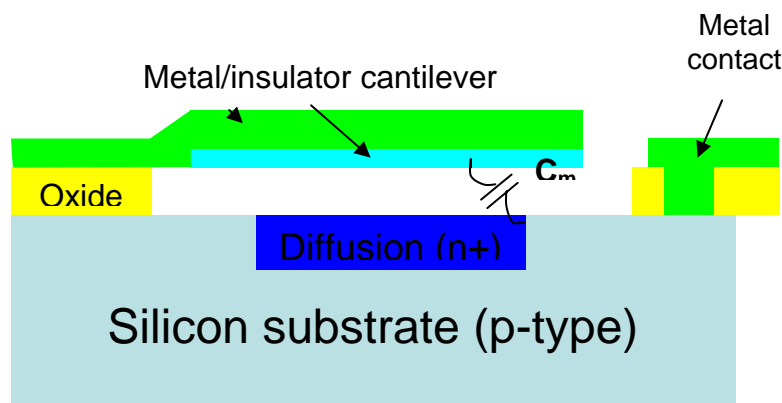
Figure 12 A MEMS-based cantilever acoustic sensor.

The cantilever is designed (dimensionally) to resonate in the acoustic range of interest. Changing the dimensions of the cantilever changes the band of acoustic sensitivities. An array of sensors, with different resonant bands, is possible. Such an array can be designed to give maximal coverage of the frequencies of interest. These can be combined to provide a faithful electronic reproduction of the signal. The major difficulty in fabrication is controlling cantilever residual stresses.

The microphone can be "read out" in a number of ways. The capacitor plates can be pre-charged and a voltage signal can be read through a low-input capacitance differential amplifier front-end, or the change in capacitance on vibration can be directly sensed by a capacitive bridge. We can integrate both approaches on a single piece of silicon.

In recent years, considerable progress has been made in the area of surface immobilization of macromolecules. Typically, a monomolecular pre-cursor (such as siloxane) is attached to a

semiconducting surface. This forms a self-assembled, dense-packed film on the semiconductor surface. Head-groups can be attached to the siloxane. These head-groups react with target analytes, fixing them on the semiconducting surface. The analytes may be toxic gases (like sarin) or anthrax spores.

The attached macromolecules themselves may have a net charge. Even if the analyte molecules have no charge, they affect the thickness of the Helmholtz double layer (solution space charge) above the attached molecule when the system is submerged in water. If these films are used as the 'gate-dielectric" of an otherwise gateless field-effect transistor (FET), the underlying charge channel in this will be effected either by the net charge of the film or by the changing coupling capacitance between the channel and the solution potential. A schematic of an FET used to sense such attachments, and fabricated and tested examples for DNA attachment experiments in Figure 13[22-25].
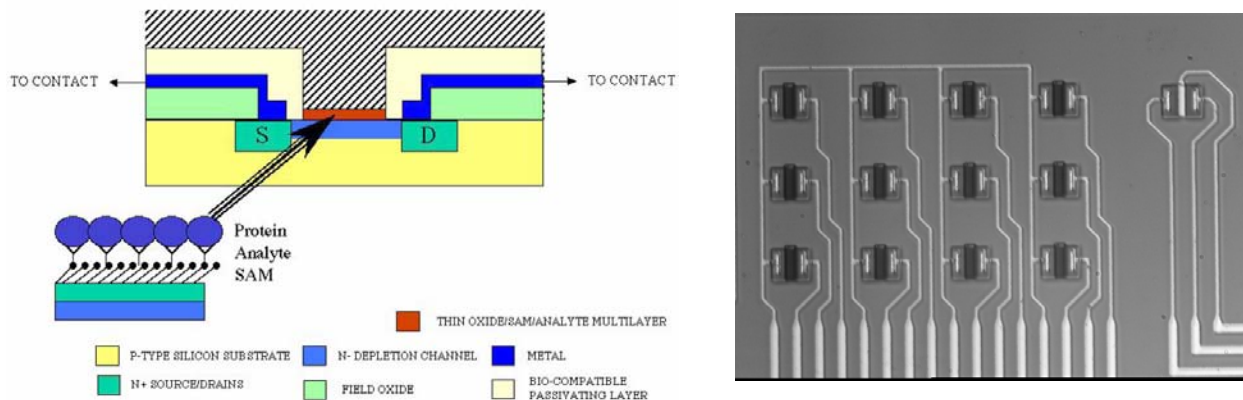

Figure 13 Schematic of a chemFET and a fabricated chemFET array.

Both of these types of sensors will be incorporated into our distributed sensor network.

8. Digital Design

In our initial prototype, the digital processing engine will be a low power microcontroller, which will integrate functionality for monitoring and adaptation of node behavior with the data processing tasks to be performed at each node. Here we will build on preliminary embedded software implementations of TDMA protocols that we have completed. These include an implementation of an epoch algorithm for inter-node time frame synchronization and a slot allocation algorithm for inter-node collision avoidance. These implementations have been developed using the Texas Instruments MSP430 family of "ultra-low-power" microcontrollers. Components in this family offer high programmability with active-mode power consumption below 500 microwatts when operating near the lower limit of their voltage range. In addition, these components provide integrated A/D and D/A converters, and processor shut-down modes in which power consumption is reduced to amounts as low as a few tenths of a microwatt during idle periods. Experimentation with microcontrollers such as these in our initial prototype will focus our digital design techniques on the performance and memory constraints that are critical to achieving the desired levels of miniaturization and low power consumption.

In parallel with the development of our initial prototype, we will explore approaches for achieving the miniaturization and 3D integration capabilities that will be targeted by our subsequent prototype designs. This will involve developing our own streamlined implementation of the off-the-shelf microcontroller used in the initial prototype. The objective of this implementation will be to maintain at least some of the programmability of the original microcontroller component, but achieve 3D integration capability by designing and implementing the integrated circuit ourselves. Mechanisms for power reduction in the implementation will center around removal of unnecessary or underutilized processor/chip features (e.g., instructions, peripherals, memory capacity, and addressing modes), and careful application of hardware/software co-design to migrate critical functionality into custom hardware implementation.

9. Energy-Efficient Distributed Algorithms for Networking and Fusion

As part of this project we plan to develop implement and optimize energy-efficient algorithms for data processing, communication, and networking that will allow distributed computation of fusion tasks over large-scale networks of randomly distributed sensors. In particular, the fusion objectives of interest in this project comprise methods for event detection, classification, and localization, as well as localized post-detection processing and tracking.

Several important constraints invariably arise in designing resource-efficient physical- and network-layer algorithms for performing fusion tasks over large-scale sensor networks. Limitations in sensor battery power and computing resources place tight constraints in the rate and form of information that can be exchanged among sensors and the type of on-sensor processing that can be performed. Equally important, changes in the network topology and size, due to node mobility node failures and battery outage, make global knowledge of the changing network topology impractical. As a result, ad-hoc networking and decentralized approaches to fusion are becoming increasingly preferable over their hierarchical centralized counterparts for such large-scale sensor networks.

9. 1. Algorithms and Computational Limitations in Wireless Networks

Some of our work in applications of wireless sensors has focused on detection and estimation problems for objects moving in a sensor network[26]. Under the assumptions that the sensors have limited communications capability and power, we have developed algorithms that can detect and track a moving object, sending a minimum of information. The algorithms have been developed and tested largely in an experimental setting. We have been able to identify individuals by the sound of their footsteps using modified speech processing algorithms distributed over a sensor network. We have been able to detect and count individuals moving past a sensor network boundary. The latter algorithm requires transmission of sensor ID, and two real numbers at the time of a detection event. Time synchronization is not required. The success of this algorithm indicates the possibility of solving interesting problems with very limited data communications, and therefore, limited expenditures of power. The research proposed here will attempt to discover the limits imposed on successful detection and classification by communications, power, and computational constraints.

In a series of papers, P.R. Kumar and his colleagues[27] have established bounds for the communication bandwidth of multi-hop sensor networks, and have, indeed, laid the foundations for an information theory for sensor networks. Among other accomplishments, they have computed the "capacity" of a sensor network framed in terms of the capacity to transport bits of data over unit distance; e.g., bit-meters/sec. A sensor network with n-elements, has transport capacity like $O\left(\sqrt{n}\right)$, and so, each sensor has an allocated capacity like $O\left(\dfrac{1}{\sqrt{n}}\right)$. From this one can conclude that only small ad hoc networks, or networks supporting mainly nearest neighbor communications, are feasible. In addition, if data is transmitted by multi-hop links from sensor to sensor to sink nodes, then the sensors near the sink nodes have to carry much more traffic than nodes far away. It follows that the power consumption of the sink node neighbors determines the lifetime of the network.

Recognizing these kinds of fundamental limitations and taking into account the advances in hardware technology developed by other team members, in the near term we propose to solve the following problems: (1) Quantify the tradeoffs between probability of detection and classification of anomalies with network power consumption (lifetime) and transport capacity; (2) Determine the "optimal" (stochastic) distribution of sensor nodes and sink nodes to maximize network performance; (3) Determine optimal distributed network management and control strategies to maximize performance and lifetime.

9.2. Physical-Layer Algorithms and Multiple-Access Protocols

The physical and network layer algorithms we plan to develop will provide nodes with the capability to communicate and fuse information across large-scale networks. In particular, based on its

limited transmit power, each node will be able to establish direct bidirectional communication with a small subset of neighboring nodes and will thus be aware them and their IDs.

In addition to the hardware-dictated ASK/FSK modulators/demodulators, the point-to-point communication algorithms we plan to employ will include source and channel coding modules. We plan to employ rate-compatible punctured convolutional (RCPC) channel codes [33], due to their low encoding/decoding complexity, and their inherent provision for different levels of error protection to different information bits. We will also leverage our prior experience with low-complexity source encoding algorithms ([28,29]) in order to implement energy-efficient systems for encoding and communicating the analog measurements in digital form over the wireless channel.

Taking into account the inherent need for ad-hoc networking, we plan to employ TDMA protocols that will be established via local negotiations over the ad hoc network. The choice of TDMA as opposed to CDMA is motivated by its multiuser interference characteristics in the large-scale networks of interest. Although CDMA can provide advantages over TDMA in cellular systems, including inherent statistical multiplexing gains and graceful increase in multiuser-interference levels as the number of users increases beyond system capacity, its implementation in the ad-hoc networks of interest imposes significant challenges. In particular, to achieve manageable multiuser interference levels in asynchronous communication settings, CDMA systems require power control at every receiving node, to set all received signals at similar power levels. Although in cellular settings this is, in principle, straightforward, in the ad-hoc networks of interest there are typically multiple conflicting requests for transmit power adjustment, as each node's data are received by multiple nodes. In contrast, TDMA systems can manage multiuser-interference by employing guard bands, i.e. without needing power control.

In order to address the inherent need for scalable fault-tolerant networking that can accommodate changes in the network topology, there is a need for distributed TDMA protocol assignments. Due to the nature of ad-hoc networking, there is also a need for spatial reuse and reallocation of TDMA time slots so as to avoid interference between transmitting nodes, while maintaining acceptable baud-rates. As part of our preliminary investigation, we have constructed distributed algorithms for establishing non-interfering slot assignments based on local network negotiations. At the initialization stage, the nodes in the network select their own time slot assignments asynchronously and independently. The protocol assignment algorithms we have constructed are local, in the sense that each node adjusts its own time slot and slot-timing reference based on observation of the signaling and requests from neighboring nodes. Provided the baud-rate is low enough, these local negotiation algorithms can provide non-interfering TDMA slot assignments. We remark that, as bandwidth efficiency may not be as important as efficient battery-power use, congestions can be always alleviated at the expense of effective baud-rates (and thus delays in fusion) by increasing the number of TDMA slots in a frame. This approach is also attractive for generating an initial operational network configuration in large-scale networks of randomly dispersed sensors. Indeed, using a TDMA allocation table with a large enough number of time-access slots (with respect to the reuse distance) can allow the nodes to self-organize and obtain an initial locally constructed multiple-access slot allocation, which can then be followed by slot reassignment at increased network baud rates. We plan to optimize these algorithms and develop extensions that also address, via local protocol adjustments, the time-varying aspects of the topologies that arise in large-scale wireless sensor networks.

9.3. Ad-hoc Networking and Fusion

The integrated networking/fusion algorithms we plan to implement will allow distributed implementation of various tasks, including event detection, localization, and tracking. Performing local processing and fusion prior to relaying data across the network can provide substantial data compression gains and as a result, savings in both power and bandwidth usage. A promising approach to performing distributed fusion involves viewing each node in the network as a local fusion host and developing integrated routing/fusion algorithms such that the fusion objectives are eventually made available to every node in the network.

In [30,32] we developed methods for local processing and fusion that allow distributed computation of various global computations, including weighted averages of functions of the node data, over ad-hoc networks with arbitrary topologies. These low-complexity algorithms are locally constructed (all that is required is that each node be aware of the subset of neighboring nodes with which it can establish bidirectional communication) and locally optimized to provide at each node fast-converging approximations to the desired global computation. They are also inherently distributed, scalable, fault tolerant, readily adaptable to changes in network topology, and suggest an optimized transmit power range for each node. As shown in [32], to optimize the total transmit power used in performing such distributed computations in networks of randomly placed nodes, the transmit power per node should be chosen so that each node can communicate, on average, with about 10 other nodes. We plan to leverage our experience in this area and in developing optimized DSP implementations to implement and optimize such distributed computation algorithms and investigate the trade-offs between processing and communication power and fusion quality over the networks that we plan to deploy.

The algorithms in [30,32] can be used to perform more complex fusion tasks, which can be decomposed into sets of weighted average computations of local functions of the sensor data. These include a number of source detection and localization algorithms. For instance, in [31], we develop distributed algorithms for source localization via an ad-hoc network of acoustic sensors, whereby each node knows its own location and has (power-based) measurements that provide relative-range information between the node and the source. In the early project stages, involving small-scale networks, we plan to build on our prior experience with such distributed source localization algorithms [31] to develop hardware implementations for sensor-location aided event localization and tracking.

At later stages of the project, we plan to consider problems of power-efficient tracking of events through the sensor field. We plan to focus on developing and optimizing *localized* fusion algorithms, whereby only a (potentially time-varying) subset of nodes in proximity to the event of interest have relevant measurements for fusion. Our proposed approach consists of methods for localized tracking via a sequence of active computation subnetworks. In [34], we have developed such preliminary algorithms that can perform distributed tracking of a moving source by forming such sequences of active subnetworks. To obtain a snapshot estimate, a computation network is formed by a one-step broadcasting to determine at each node the subset of neighboring nodes contributing to this snapshot computation. The computation algorithms are then set up by locally refining the given active network's base distributed computation modules. As an example, tracking a source with this approach with active subnetworks of approximately 50 nodes (and with optimized transmit node power) requires about 15 communication steps to set up and perform the active subnetwork computation. Considering, in addition, the distributed computation of a weighted least squares localization algorithm, 5 parallel computations per step are required for localization. Assuming also a 2-byte precision in the computation (and ignoring the multiple access protocol control information overhead), snapshot computations of this algorithm can be performed at a rate $R/(150N)$ snapshot estimates per second, where R denotes the point-to-point bit-rate, and N denotes the number of slots in the multiple access protocol. In this context we plan to leverage our prior experience involving methods for adaptively forming active subnetworks with optimized distributed computation algorithms with via local refinements [34], to develop and implement energy-efficient algorithms for distributed event tracking in the large-scale networks we plan to deploy.

9.4. Network Configuration and Optimization

The algorithms described in the previous sections for the physical and network layers and integrated networking/fusion will be most effective when they are configured and mapped to hardware/software realizations in an integrated manner that takes into account the physical constraints imposed by our 3-D integration technology and the desired levels of accuracy, robustness, and performance in the application. The parameters and design evaluation metrics underlying these setup issues form a vast and complex design space, whose efficient exploration will lead to significant power savings, thereby extending the lifetime and reliability of the network. Therefore, careful and integrated optimization of these parameters based on data from design-time models and simulations will be useful during the system construction/initialization process. Furthermore, feedback from this type of

optimization will be helpful in refining the algorithms themselves so that they are more tightly matched to the embedded processors and more effectively parameterized to exploit the underlying trade-offs and to further facilitate system-level optimization.

Central aspects of the underlying design space here include the type and amount of processing resources to allocate to each of the algorithms; configuration of hybrid compile-time/run-time scheduling of the mapped hardware and software; policies to apply for dynamic voltage and frequency scaling on the processing platform; and system-level parameters of the transceiver circuitry, particularly the PLL or the microcontroller. Algorithms and software tools for efficiently exploring this design space will allow us to more fully exploit the potential of our device, fabrication, and communications technology, and will allow designers to use this technology more easily.

To address such system-level optimization issues, we will explore the use of evolutionary algorithms, which provide a probabilistic search methodology for complex, non-linear optimization problems[35]. In this work, we will leverage our previous experience in architecting evolutionary algorithms for synthesis and optimization of embedded hardware and software (e.g., see [36,39]). The novelty of our effort here will be in its emphasis on integrated optimization across the configuration spaces of sensor network algorithms, hardware, and software; and its building on efficient new techniques for hybrid optimization, multidimensional search, and evolutionary algorithms that we have been developing in recent years at the University of Maryland. Indeed, the complexity and importance of ultra low power sensor network applications make this family of applications an especially useful driver for development of hybrid and multi-dimensional optimization technology.

In evolutionary algorithms, candidate solutions to an optimization problem are encoded as binary strings based on a "genetic representation" function. This "population" of binary strings is then adapted probabilistically through operations such as "recombination" (combining attributes from two existing binary strings to form a new binary string) and "mutation" (making small perturbations to an existing binary string). As these adaptations occur, the modified population is filtered with a strong bias toward retaining those binary strings that correspond to the most effective candidate solutions. This process of probabilistic adaptation and filtering continues through numerous iterations, and the best candidate solution in the resulting final population is chosen as the output. When the key evolutionary algorithm building blocks, such as the genetic representation, recombination and mutation operations and the filtering mechanism, are designed well, evolutionary algorithms have been shown to be effective across a wide variety of complex optimization problems (e.g., see [39]).

When implementing applications such as the proposed location applications (described further in Section 9.1) using our small sensor network technology, we must consider latency, throughput, and peak and average power consumption. The tools for network configuration that we explore in this work will therefore require complex, multi-dimensional optimization methods. Complex optimization methods are often enhanced by efficient local search algorithms, which exploit the structure of specific sub-problems and refine arbitrary points in a search space into better solutions. In many cases, these local search algorithms can be parameterized so as to trade off time or space complexity for optimization accuracy.

Furthermore, these trade-offs may span orders of magnitude in each of these dimensions (complexity and accuracy). We have demonstrated the importance of parameterized local search algorithms (PLSAs) in embedded systems synthesis, and of carefully managing their associated run-time/accuracy trade- offs when PLSAs are used to refine points generated by supervising global search algorithms. We have developed a novel framework called "simulated heating" for this purpose of run-time/accuracy management[37]. We have developed both static and dynamic trade-off management strategies for our simulated heating framework, and have evaluated these techniques on a variety of practical optimization problems with very different structure. We have shown that in the context of a fixed optimization time budget, simulated heating better utilizes the time resources and outperforms conventional fixed parameter search methods.

We will build on our experience with PLSA-based optimization methodologies and evolutionary algorithms to explore methods for developing, parameterizing, and integrating efficient local search techniques for synthesis of efficient sensor network configurations. This will include exploring PLSA

formulations for key algorithm, hardware and software configuration subproblems. Leveraging our previous work on evolutionary algorithms for multi-dimensional optimization of signal processing software [38], we will also explore new global search strategies that work effectively for the PLSAs, design constraints, and parameter spaces associated with our proposed class of millimeter-sized sensor networks.

Complex optimization techniques of this nature will benefit from powerful computational resources. For this purpose, we will utilize the high performance cluster of integrated graphics processing units (GPUs) and high-speed, general-purpose central processing units (CPUs) that has been awarded to us through an NSF Research Infrastructure grant that is being led by collaborators in the University of Maryland Computer Science Department. GPUs form an important class of domain-specific processors that are finding applications in graphics as well as other application areas that involve regular, vector-oriented operations. These kinds of operations are in fact characteristic of many evolutionary algorithms. The use of GPUs for greatly improving the performance of evolutionary algorithms (specifically, a subclass of evolutionary algorithms called genetic algorithms) has been demonstrated previously by applying GPUs to a standard evolutionary algorithm for the well-known 3-satisfiability problem. Using our high-performance GPU/CPU cluster, we will explore techniques for efficiently parallelizing the evolutionary techniques that we develop in this proposed project. In this effort we will investigate techniques for designing evolutionary algorithms and global search/PLSA integration methods that map especially efficiently to the heterogeneous capabilities of our cluster. We will evaluate the quality of configurations synthesized by our optimization techniques initially by simulation and in later years by deployment on the small sensor network test-bed that will be developed concurrently in this project.

10. Education

We plan to develop two new education projects that are synergistic with the proposed smart sensor program. One project is the development of a new Capstone Design class and the other is to have an undergraduate Gemstone education/research group arranged around the smart sensor network concept.

The Capstone Design class, which we will call "Wireless Smart Sensor Networks," is ideal for this project because it is interdisciplinary, and it draws on many of the areas that students are expected to study during their undergraduate years. If you consider a wireless sensor network, it has components from almost every field in electrical and computer engineering. Smart sensors are actually small smart systems. The network itself is an application of communication theory and electromagnetics. In the class we plan to have students build their own elementary smart sensor network. This will require students to apply their skills in electronic design, microprocessor programming, digital and analog communication techniques, as well as electromagnetic theory. Topics will include communication protocols, power management and lifetime prediction, and distributed algorithms for detection and estimation. Finally, since engineering ethics is part of the capstone design curriculum at the University of Maryland, becoming aware of the ethical considerations surrounding smart sensor networks will force both students and faculty to consider how to best deal with these issues.

We will form a Gemstone Team to pursue a project involving wireless sensor networks. The UMD Gemstone Program (www.gemstone.umd.edu) brings together talented undergraduates from all disciplines in the University to work on a focused project over a 3 year period. A typical team has 10 to 15 members. The teams conduct research and write a thesis, which is presented to a panel of experts in the senior year. This year we have a senior team of 12 students that has been working on technology for First Responders, including sensors for measuring firefighter health and the surrounding environment, and sensors for detecting police officer actions (weapons use). This Team has produced novel technologies in each of these areas. If this project is funded, we will form one or more Gemstone teams to pursue research on applications of wireless sensors in health monitoring, first responder safety, etc.