THE APPLICATION OF REMOTE SENSOR TECHNOLOGY TO ASSIST THE RECOVERY OF RARE AND ENDANGERED SPECIES

Edoardo S. Biagioni
K. W. Bridges

UNIVERSITY OF HAWAI'I AT MĀNOA, USA

We describe a wireless sensor network designed for the long-term study of rare and endangered species of plants. We wish to monitor plants and their environment via high-resolution cameras and temperature, humidity, rainfall, wind, and solar radiation sensors. Our units must be "invisible" (camouflaged), very low energy, and must allow distributed local computation. Data rates are 1 to 100 bytes/second per node, but networks can be large – an early prototype had 60 nodes. Failures are expensive and we must exploit redundancy whenever possible. Nodes are stationary but for energy reasons may decline to participate in transmissions.

We have designed two wireless routing protocols that satisfy these constraints. Multipath On-demand Routing (MOR) computes multiple optimal routes to avoid depleting the energy at any given node. Geometric Routing scales to large networks, relying on Geographic Routing when possible and on selected global information otherwise. We have simulated and are implementing both protocols.

1 General Problem Design

A rich set of environmental information is a fundamental resource needed to help preserve federally listed species (U.S. Endangered Species Act, 1973). This includes long time series measurements of the temperature, humidity, rainfall, wind and solar radiation in the rare species' habitat. In addition, it is useful to have the same information for nearby areas that do not have the species so that it is possible to make inferences about the climate as a factor determining the species' distribution.

Monitoring species performance, particularly the phenological events such as the periods of active growth, flowering, seed-set and the like, is equally critical to understanding why a species is rare and to evaluate possible remedial actions.

The International Journal of High Performance Computing Applications, Volume 16, No. 3, Fall 2002, pp. 000-000

Setting up a comprehensive environmental measurement network and making regular observations of the species is not only expensive, it is difficult to accomplish with minimal impact on the species that are being assisted. For example, frequent visits can be a source of habitat modification that might place the species being studied at further risk.

The PODS project (http://www.pods.hawaii.edu) has designed a network that addresses the constraints of the biology of typical rare plant species and the habitats they occupy. Our initial test location is in Hawaii Volcanoes National Park (Hawaii Island, Hawaii, USA). We expect that this design is quite general and we will test it in a number of extreme environmental situations.

The choice of a primary study area was governed, in part, by the presence of steep environmental gradients within the park. For example, there is an apparent rainfall gradient over the 6 km distance from the rain forest near Park Headquarters to the near desert-like conditions at South-West Rift Zone site with the threatened *Silene hawaiiensis*. Other apparent differences include the amount of cloud cover and wind. We do not know the magnitude of these differences because there are no environmental measurements available for the South-West Rift Zone.

2 Network Configuration

Each unit has a computer, transmitter and receiver in addition to the environmental sensors and, in some cases, a high-resolution digital camera (Figure 1). Each sensor unit collects data and moves data, both its own and that from other units, toward an Internet node.

There are two general configurations of sensor units. The sensors are dense in the area in which the rare species occurs so that a full range of microhabitats is fully sampled. The other configuration is less dense and it serves a dual purpose of monitoring the environment adjacent to the area with the rare plants and of providing a communications pathway back to an Internet node (Figure 2). The routing of this communications pathway is determined by both the need for environmental information and to be a relatively direct path between the intensive area and the Internet node.

The spacing of the sensor units provides redundant communication pathways within each of these deployment areas. This allows continuing communication even in the event that a node should fail.

3 Major Design Constraints and Opportunities

Research in the national park limits the type of equipment that can be deployed. At the least, it must not harm

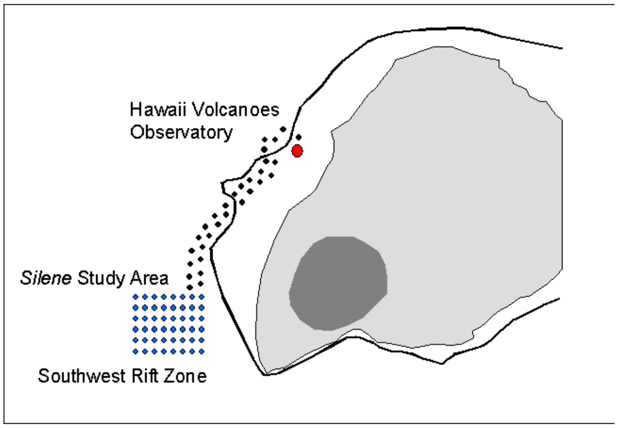

Fig. 1

any species. It is also a requirement that the equipment not disturb the visitor experience. In the case of the South-West Rift Zone, this is a particular concern since it is an area that receives a large number of visitors. The South-West Rift Zone is an open landscape with large flat areas that are interrupted by shallow depressions. These depressions are the "cracks" that form as the magma forces the Kilauea Volcano away from Mauna Loa, the adjacent volcano that is the most massive mountain on Earth. They are the geological attractions that receive the visitor's attention. There is little vegetation in the flat areas. It is mainly a few widely scattered small shrubs, including the threatened species, *Silene hawaiiensis*. The surface soil consists of an ash deposit that has been consolidated into a hard layer by many years of acidic rainfall. The area also has some rocks, many with a diameter of 0.2 to 0.5 m, that were ejected from the nearby Halema'uma'u crater.

There is little that can be done to hide instrumentation near the South-West Rift Zone. We chose the rocks as a candidate since they are least likely to draw attention. One of our design goals is to produce "fake rocks" that appear similar to those naturally occurring in the area. These are hollow so that they can conceal the sensors, computer, radio and a power source.

The closest Internet connection is approximately 2 km from the South-West Rift Zone. It is located in the Hawaii Volcanoes Observatory, a unit of the U.S. Geological Survey. The Observatory is located at a slightly higher elevation and there is rough, undulating topography en route. There is also a vegetation change, likely the result of slightly more rainfall, with scattered small 'ohi'a trees (*Metrosideros polymorpha*) and larger and more abundant shrubs. There are no particularly outstanding visitor attractions along this route, aside from the distant views of Halema'uma'u Crater and the slopes of Mauna Loa.

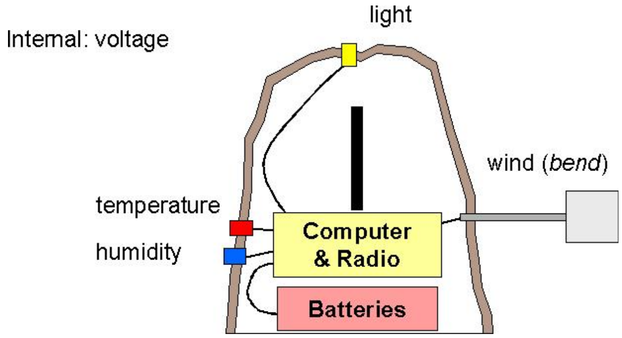

Fig. 2

The availability of small trees expands the possibilities for hiding the instrumentation. We chose short, hollow structures that appear to be tree branches as possible containers. These can be placed in the trees at some distance from the highway and on the opposite side of the road from the usual vista enjoyed by the park visitors.

Other design considerations include a desire to be able to remotely monitor and administer the network.

3.1 SENSOR CONTAINER DESIGN AND CONSTRUCTION

The containers needed to be "not obvious" and able to conceal the equipment. This meant that they had to be as transparent to radio signals as possible.

The "fake rocks" (hereafter called "rocks") were made from the micro-fiber filler that is used to repair dents in automobiles (often called by the trade name of "Bondo"). This material was selected after experimentation with several other casting compounds. The mold was made from several layers of latex and gauze. This thin mold holds the textured surface of actual rocks very well. The shape of the latex mold was maintained by making a plaster of Paris outer mold (split into two parts to make it easy to remove the casting). The filler is an epoxy substance that cures when hardener is added. A colored pigment was added to the filler prior to mixing in the hardener. The rock is cured in about 10 minutes and can be removed from the mold. Considerable heat is produced during the curing and this limits the useful life of a latex mold to about a half-dozen castings. The final step is the trimming of the bottom of the rock so that it sits flat on the ground.

The rock is an open half-shell. Efforts to seal the rock were sufficiently challenging that we have instead chosen to use watertight bags to seal the equipment that is placed inside the rock. These are "bead closure" (Ziploc) bags. A small hole in the bag provides an opening for cables. The opening is closed around the cables with a cable tie.

This arrangement makes efficient use of the space and allows the bags to be opened if the equipment needs to be manipulated. One bag is used for the computer and radio and another for the battery power supply.

The “tree branch” containers are built from approximately 30 cm sections of 6 inch (15 cm diameter) PVC pipe. Testing caps (which are thin and fit inside the end of the pipe) are used to seal the top end of the container. A micro-fiber filler cap is constructed to fill in the testing cap and give it a more natural appearance. The bottom end of the pipe is left open. A 1 inch thick Styrofoam circle is fit below the testing cap to act as insulation and to provide some water resistance. Another such circle is used at the bottom to secure the battery pack and to hold the computer and radio in place. The white PVC pipe is disguised by gluing black-and-white laser printed pictures of actual tree bark to the outside of the pipe. Although a number of color tests were done to simulate the appearance of the actual trees, the trees in this area appear closer to black and white than any of the other colors that we tried. This paper covering is then sprayed with a waterproof spray. Holes drilled through the base of the pipe permit the use of long cable ties to attach these containers to the trees. Placement of these fake “dead tree branches” at the place where trees’ branches fork makes them very inconspicuous.

Sensors are mounted in most cases by drilling through the container and using hot glue to secure the sensor and make a watertight seal.

3.2 SENSOR SELECTION

The most important environmental parameters to monitor at the South-West Rift Zone are air temperature, light, wind, relative humidity and rainfall. Sensors for all of these except rainfall are installed on the rocks, although not all rocks have all of the complete set of sensors.

At this stage in our development, we are interested in getting instruments into the field that let us generally monitor many things at many places. This is in contrast to the more usual approach where just a few parameters are measured with high precision at a few locations. This is due, in part, to the fact that this area has steep environmental gradients that we feel are important to measure. Perhaps more fundamentally, however, we need to gain experience handling a large data flow and discovering the problems of maintaining a large network of sensors. As a result, we have chosen as simple a set of sensors as possible. All of these easily interface with the TephraNet platform (see Section 4.1).

Temperature is measured with a thermistor (DigiKey). It is located low on the side of the rocks, if possible under a slight projection. In addition, the rocks are deployed so that these sensors will be on the north side of the rock. This generally avoids heating by direct sunlight. The thermistors on the tree branches are inside the bottom of the container. In this location they are in the shade and receive adequate air circulation.

The light sensors are photoresistors (Image Company). They are placed horizontally at the top of the rock and similarly on the top of the lip of the tree branches. The goal of these sensors is more to detect the presence of clouds than to measure the solar radiation. As a result, differences in calibration and mounting angles are not important. We are looking at the relative amount of light received by an individual sensor.

Many different wind sensors were considered. Cup anemometers and other spinning designs were rejected because of the difficulty in hiding them. Although a Pitot design can be hidden, we were not able to procure or manufacture one to fit our needs in time. Instead, we made a simple sensor from a bend-sensitive device. This is a flexible piezoelectric strip (Image Company) that is 5 by 100 mm and that changes resistance when it is bent. We added a small flap to the end to provide a larger surface on which the wind would push so that we would have better sensitivity in the expected range of the wind. This was roughly calibrated in a simple wind tunnel with an accurate anemometer. Since this sensor is slightly obvious, it is not placed on the rocks that will be near park visitors. It is located on the side of the rock near ground level. Besides the relatively poor resolution of this sensor, it only measures the wind from a single direction. As a result, we will depend on the network of these sensors to provide general information about the wind speed and direction.

Relative humidity sensors are often included in the minimal set of environmental measurements. While we recognize their value, we have not added these sensors to our basic sensor package. The need for these sensors is being reviewed and the rock design includes a place for them to be mounted. If they are to be included, there are several low-cost sensors that are available.

Rain sensors were not included at this developmental stage for several reasons. Tipping rain gauges, the standard type of rain sensor used in this sort of situation, are large and their collection opening is very obvious. We are experimenting with an alternative design that provides information on when there is rain, rather than the amount of rain that is collected. This alternative consists of conductive materials that are separated by narrow spaces. Water falling on this sensor changes its conductivity. The type of materials used, the spacing of the conductors and the angle at which the sensor is mounted allow this design to detect the period of the precipitation and cover a range of moisture sources, such as differentiating between fog and rainfall. This type of sensor is not equivalent to a traditional rain gauge, but it is compatible with the requirement for being not obvious and will record data that is useful to the problem of recovering rare species.

4 Communications for Sensor Networks

The most significant factors in the design of the communication system for distributed sensor networks are ease of deployment and reliability. Ease of deployment requires wireless communication and low weight. Reliability requires continued function even if some of the communication relays are lost.

Wireless communication is available commercially in several systems, or can be designed in-house using the ISM (Industrial, Scientific, and Medical) radio bands. Commercially available protocols include IEEE 802.11, which is a wireless LAN standard, and Bluetooth, which is designed as a replacement or upgrade for IrDA but may be extended to larger distances. Both of these work in the 2.4 GHz ISM band. Short hops as provided by Bluetooth and 802.11, in the tens to low hundreds of meters, are suitable for environmental sensing since they allow us to sense conditions at intermediate points between areas of interest, providing useful information about the areas where the endangered plants we are studying are not present. This information complements the information we gather in the areas where the endangered plants thrive and allow us to infer the effects of the environment on the plants' survival.

We first give a brief overview of a wireless communications network designed and built by a team at the MIT Media Lab with whom we have been co-operating, and follow with a description of the communication network that we are currently developing.

4.1 TEPHRANET

We are co-operating with Michael Hawley's MIT Media Lab research team that has designed and built systems – Tephranet – to incorporate into the environmentally inconspicuous containers ("rocks" and "tree branches") described above. The Tephranet system consists of a radio and an embedded, battery-powered computer capable of reading the sensors. The radio works in the 900 MHz band and, based on preliminary measurements, appears to have a range of approximately 30 meters when placed 10 cm above the ground. At a height of 2 meters above the ground, the range increases to about 100 meters. In the intensive study area, we plan to deploy "rocks" within 30 meters of each other. The intensive study area is approximately as wide as it is long, and we fully expect each rock to be able to communicate, on average, with four to six neighbors, so that even if one or a few neighbors should go down, functioning sensor pods can still communicate with each other. Figure 2 shows our planned deployment.

To reach our Internet connection 2 km away from the nearest edge of the intensive study area, we plan to place units in small trees to raise them to at least 2 meters off the ground. As a result, we can plan to space units farther apart and overall use fewer units in this communication pathway connecting the intensive study area to our base station. Since we wish to avoid having a single point of failure cripple the entire network, each node in this communication pathway is deployed so as to be able to communicate with at least two other nodes in the forward direction and two in the reverse direction. This means we expect to place a unit approximately every 50 m, and we need 40 units to cover the 2 km distance. Overall, we are deploying about 100 units. The 40 in the communication pathway should be able to communicate end-to-end in about 20 hops, and the intensive area, of a size 8 by 8 hops, should have an approximate diameter of 10 hops, so that the overall network diameter may exceed 30 hops. We believe this is a relatively large sensor network compared to other sensor networks being deployed at the present time.

One concern for deployment is the weight of the pods. The enclosures, computers, and sensors weigh relatively little, and much of the weight of a pod is in the batteries. The Tephranet systems are designed to be off most of the time, giving the most lifetime for a given weight of batteries. Every 10 minutes, the units automatically turn on for a short time, record and communicate the values of the sensors to the base station, and turn themselves back off. We estimate that a Tephranet unit can perform this function for about 6 weeks powered only by 4 "C"-size batteries.

The protocol used to route and communicate is called GRaD and allows the determination of connectivity and routing information as well as the transmission of sensor data and other statistics used for network management. The protocol also allows the different systems to synchronize their clocks, which is important so they can agree on when communication should occur.

4.2 THE PODS NETWORK

While the Tephranet systems are sufficient for many of the functions that we envision, we have also embarked on the design of a "next generation" computer and radio communications system. Some of the additional features we would like to see in such a network include the use of standard networking protocols (including support for TCP/IP) and of accepted wireless networking standards, allowing the purchase of commercial communications equipment.

One obstacle to using commercial equipment is the current lack of a wireless routing standard for *ad hoc* wireless networks (802.11 provides a standard for *ad* *hoc* wireless networks, but every unit must be in range of every other unit). When first deployed, a network of wireless sensors must engage in discovery to learn about the nodes it can directly communicate with. Once the directly reachable nodes have been discovered, the network as a whole must execute a routing protocol so that each sender learns how to reach each of its destinations. Since we wish to have flexibility about the roles of sender and receiver, the routing protocol must provide each node with information on how to reach all the other nodes. Moreover, during deployment more nodes are added gradually to the network, and nodes may occasionally move. After deployment is complete, nodes may lose power or drop out of the network for other reason. The routing must dynamically adapt to these conditions, but should require very little power (i.e. very few transmissions) as long as conditions are stable. Finally, we can envision having a sufficiently redundant network that not all nodes are needed for connectivity on each transmission. If each node has enough information to avoid participating in a communication, it might save power either by not transmitting the data, or even by remaining turned off for the entire time.

There are some existing routing protocols for wireless *ad hoc* networks. Above we mentioned GRaD, used in Tephranet. A similar protocol is known as "diffusion." Both these protocols are optimized for sending data to one or a few base stations, rather than for generalized routing. For our application, we eventually envision pods communicating with each other, both within the local area, to compare measurements, and outside the local area, to track and map events of interest, for example rainfall or temperature.

There are protocols, such as DSR, Dynamic Source Routing (Johnson and Maltz, 1996) and DSDV, Dynamic Sequenced Distance Vector (Perkins and Bhagwat, 1994), and others (Broch et al., 1998, and especially Jones et al., 2001). These routing protocols are suitable for generalized exchanges but have other limitations. DSR in particular must include a complete route in every packet that is sent, and therefore would require unreasonably large packets for networks, such as ours, which have moderate or large diameter. Geographic routing protocols, where a packet is sent to the neighbor closest to the destination, are attractive for wireless, since there is usually good correlation between the geography and the topology of a wireless network. Unfortunately, such geographic routing protocols fail in the presence of "holes" in the geographic distribution of nodes – they are unable to prevent routing to a dead end.

In geographic routing, it is assumed that there is a way to compute the physical location of a node given its network address (Imielinski and Navas, 1996). This same assumption is made by geometric routing.

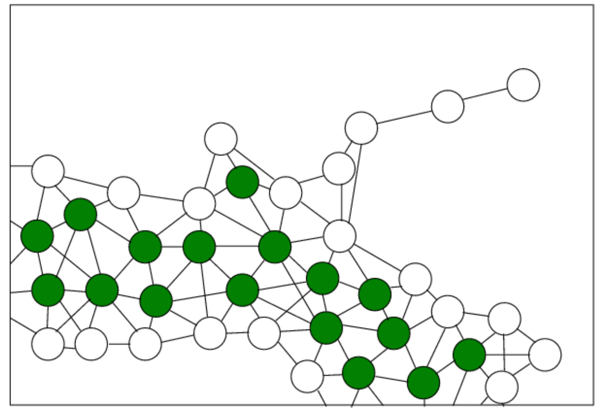

Fig. 3

4.2.1 Geometric Routing. The first step of geometric routing requires that a node determine whether it is surrounded, that is, whether geographic routing can be used, or whether it lies on an edge of the sensor network polygon. This situation is illustrated in Figure 3. The edge could be the exterior perimeter or the edge of an interior hole. Nodes that are not completely surrounded communicate with all other nodes on the same edge, so that each node on the edge builds a map of the entire edge. Nodes then communicate via geometric routing (greedy routing) whenever possible. When geographic routing fails, that is, at dead ends, data is forwarded along the shortest edge towards the destination. Data leaves an edge once it reaches the part of this edge that is nearest the destination.

To determine whether a node is surrounded, we assume that each node can broadcast a message announcing its presence and position. The nodes that receive this message are the node's *neighbors*. If a node listens to all its neighbor's announcements, over time that node can build an accurate map of its entire neighborhood. All nodes then exchange their neighborhood maps, so each node has complete information about each neighbor's neighbors. Each node sorts its neighbors according to increasing angle (bearing). Two nodes that are next to each other in the sorted list may or may not be each other's neighbors. If they are, then the area in the triangle between the node doing the analysis and the other two nodes is inside the sensor network. If the two neighbors are not each other's neighbors, the angle between them represents a gap, and the connections between this node and each of the two neighbors is part of an edge. An example of a portion of a network with edge nodes and interior nodes is shown in Figure 3. Typical sensor networks used for environmental monitoring will often have more interior nodes than shown in Figure 3, and relatively fewer edge nodes.

Geometric routing requires that complete information about each edge be distributed to every node on that edge, and hence has a relatively high initial overhead. However, once the initial information has been established, geometric routing can be very efficient. Packets are forwarded along the shortest path when no obstacles are present, and packets from different senders to different destinations naturally follow different paths, randomly equalizing the power consumption across all the network nodes. When obstacles are present, packets take the shortest way around the obstacle, then resume geographic routing. In short, the only global information that needs to be propagated is information about the edges, which for large networks is potentially much less (linear rather than quadratic) than the information about all possible destinations that, in the worst case, is needed by most other protocols.

Once information about the edge is available to all the nodes on that edge, it is at least conceivable that one could use techniques from computational geometry, for example Plane Sweep (Nievergelt and Hinrichs, 1993), to compute a route that is optimal. Finding this route may require substantial computing resources – in the extreme case, the network could look very much like a maze, and exhaustive search is the only way to find the optimal route.

4.2.2 Multi-path On-demand Routing. The second routing protocol we have developed, Multi-path On-demand Routing (MOR), uses node reachability rather than node positions. This protocol resembles somewhat the GRaD protocol used for Tephranet. When a node A wishes to communicate with a node B, it first searches its routing tables for a route to B. If one or more such routes exist, each route specifies a next hop to use to forward packets. One of these routes is selected, and the packet is forwarded to the corresponding next hop using unicast transmission. The next hop forwards the packet in turn, and node A overhears the transmission and uses it as a "passive" acknowledgement of the successful link-level transmission.

If no route is found in the routing table, node A initiates a network flood to compute a network-wide gradient rooted at A. Packets in such floods are broadcast with a random delay. Because the gradient is computed at each node, nodes closer to the root avoid retransmitting packets received from nodes that are the same distance or farther away, and the broadcast always terminates. As nodes are forwarding gradient packets, they also update their routing tables with routes to A. In each of these routes, the distance is the distance computed by the gradient, and the next hop is the node from which the gradient packet was received.

If node B is in the network, it will receive this flood and reply. The reply is unicast to each of the next-hop nodes on the routes that B has built to A. These nodes in turn forward the reply along their routes to A, meanwhile recording a path back to B. A node can forward a packet (data and routing packets alike) as long as it has a route to the destination – whenever such a packet is forwarded, a reverse route to the sender is automatically recorded.

One observation about MOR is that if there is one or a few nodes that most other nodes need to communicate with, such nodes can perform an initial network flood, creating a gradient, rooted at themselves, throughout the entire network. A base station connecting a sensor network to the Internet, for example, can do this, and establish a gradient with minimal effort. Every node that needs to communicate with such a base station then has a route to it. Also, nodes forwarding data for other nodes automatically acquire routes to these nodes, without additional overhead. We therefore expect MOR to be very parsimonious in transmission, and hence very energy efficient. Furthermore, if multiple paths to a destination are available, MOR will use all of the paths in rotation, distributing among the largest possible number of nodes the energy needed to forward packets to their destinations.

MOR is described in detail in a thesis proposal by Shu Chen, "Routing in a Wireless Sensor Network", which is available at http://www.pods.hawaii.edu/~shuc.

4.2.3 The PODS wireless sensor network. A general observation about many of the other routing protocols that have been proposed in the literature is that they have not benefited from widespread practical use, and instead have only been evaluated under simulated conditions. These conditions usually involve some sort of randomized node "density" (and sometimes motion, which is not relevant in our case). In contrast, our simulations address not only worst-case scenarios and random distributions, but also specific deployments that we have planned.

We note that, for most routing protocols, routing information only really needs to be sent when the topology changes. Some wireless routing protocols, designed to support mobile hosts, will time out and remove data that has not been refreshed within a certain period of time. It is clear, that for wireless sensor networks, route information should be kept for as long as communication is occurring, and should not be unnecessarily discarded even when a neighbor ceases communication, since that neighbor may become available again later.

One advantage of inexpensive all-to-all routing is we can run standard timing protocols such as NTP that with relatively few messages have a good likelihood of synchronizing different nodes to within a fraction of a second of each other. The power consumption of a node is approximately a linear combination of the amount of time the node is powered and the amount of data the node transmits (Feeney and Nilsson, 2001). By having an efficient routing protocol we can reduce the amount of data that a node transmits, and by having fast routes and good synchronization among nodes we can reduce the length of time a node has to be on.

Different wireless protocols take different approaches to energy conservation. Bluetooth includes extensive provisions for turning units off and on, whereas 802.11 generally requires a lengthy synchronization period to "acquire" a connection. On the other hand, 802.11 is designed to work with more sensitive receivers, which implies the senders have to send less power to cover the same distance. Ultimately we envision having some of our nodes connected with Bluetooth, perhaps the longer-range (and more power-hungry) class C Bluetooth radios, and others connected via 802.11. Since we plan to use this network to carry IP packets, we can have different data-link layer protocols in different parts of the network.

Keeping the routing protocol simple is a necessity, since simpler routing protocols can be run on less power-hungry hardware. However, it is our goal at present to maintain a lot of flexibility in our sensor platforms, allowing us to remotely upgrade software on deployed pods as well as manually trigger specific data collections, such as an intense observation period or taking pictures. This flexibility means we cannot adopt very simple and low-powered systems. We are also studying hybrid deployments where some of the nodes are higher-powered and flexible, and some nodes are lower power and more restricted in what they can do.

5 Energy Issues for Wireless Sensor Networks

In designing our sensor pods, we have encountered issues that are familiar to others who have designed wireless units. The two sides of the energy and power equations are consumption and supply. To reduce consumption, we use the lowest-power processors and radios that will suit our purpose, and keep them turned off as much as possible. This section addresses the supply side, specifically issues of power generation and energy storage in wireless sensor networks.

We are open to using a variety of sources of energy available in the environments we are studying. These include direct solar (light) energy, solar heating, and wind. Wind is usually the hardest to camouflage, since it is harvested by windmills, though we are studying the option of having holes in our "rocks" which would allow the wind to flow through and power a concealed wind turbine. Instead, we have been mostly focusing on solar power. Solar power is less available in specific environments, for example dense forest canopies, but in such environments we may be able to use bigger collectors, so the total energy available for each unit may be the same. Solar panels are also hard to camouflage.

We have begun to explore the use of Peltier effect energy generation. A Peltier unit (also known as a Seebeck or thermoelectric unit) will generate electricity if the two sides are kept at different temperatures (though the most common use of Peltier units is to pump heat from one side to the other by providing a sufficient voltage across the terminals). We have not investigated other forms of energy generation from solar heat, though they are undoubtedly possible.

For energy storage, we have been using alkaline cells (batteries) and lithium cells. With non-rechargeable batteries, the main concern is that any increase in total energy stored requires a proportional increase in the weight and bulk. According to the Energizer web page, http://data.energizer.com/, a 9-volt alkaline battery will provide almost 600 mAh, a 9-volt lithium battery about 1200 mAh.

Rechargeable batteries are useful in combination with energy generating technologies. With sufficiently low power requirements, the weight and bulk of rechargeable batteries are not an issue. Efficiency is an issue, since the amount of energy that needs to be generated to recharge the batteries increases with less efficient batteries. There are many rechargeable battery technologies of interest, including Lithium-Ion and lead-acid. For very low power levels we have found we can use capacitors, which are low in weight and 100% efficient, but relatively high in bulk. Our design goals are as follows. In the absence of energy generation, we would like a pod lifetime of at least three months, increasing to at least a year for pods deployed in hard-to-reach rugged areas. For pods that have some energy generation equipment, our goal is about a week lifetime in the absence of sunshine, with 14 hours the very minimum. Having a one-week lifetime will allow us to monitor unusual and disruptive weather phenomena, such as hurricanes, as well as simply dealing with rainy periods that effectively reduce the level of the sunshine to below that needed for full recharging.

We have also been drawn to other energy storage technologies, particularly fuel cells. At this point, we have not been able to locate fuel cells small enough for our purposes, but we are still investigating and hope that the technology will evolve quickly enough to be of use in the near future. Fuel cells are especially promising in situations where solar cells and solar heating are not valid options.

6 Data Flow and Visualization

Two types of data are collected by the sensor network and moved to the Internet. Weather data are collected every ten minutes and high-resolution images once an hour. The data repository is a server located at the University of Hawaii at Manoa (on a different island than the data collection).

The weather data are stored in a database and the images are separate files.

We expect to migrate from this standard form for storing numerical data to one that eventually will reduce the data flow. In that scheme, we will use the computer at the site of the sensors to categorize each of the values into a five level scheme. The ranges for the categories will approximately represent values that are near "normal", a small amount above or below "normal", and two more extreme categories. Such categorizing cannot be done, of course, until we have sufficient data to determine "normal" and an appropriate range for each of the categories. Further, this is not a set of fixed categories but one that corresponds to an appropriate model. For example, air temperatures are known to follow a general trajectory in which it is warmer in the day than at night. If the average shape of this curve is known, then the categories can be fitted around this average pattern. The advantage is that no data need be transmitted for the periods in which the values are within the "normal" category, other than a periodic network exchange to verify that the sensor units are functioning and communicating, and to synchronize the units' clocks. This is called the "exception reporting" scheme since the only significant reporting is done when values fall outside the expected range.

Exception reporting closely matches the type of monitoring that is interesting to the biological problem. Baseline information is developed in this scheme by the production of a model that describes the expected environmental conditions. This is reported along with the periods over which this model properly describes the environment. The other information obtained is the periods and degree of departure from the model. This is of particular interest because it is during these intervals that it is likely that significant things are happening to the environment or the organisms at the site. For example, higher than expected rainfall can result in significant erosion, or much colder than expected conditions can result in the death of small plants.

The number of sensors complicates the visualization. Typically there will be between 50 and 100 sensor modules used on a site, each sending between two and five data values. The visualization strategy follows the exception-reporting scheme and simplifies the problem. A simple display shows colored dots, one for each type of sensor at each of the sensor modules. The total anticipated number of colored dots is about 500. The color-coding shows the category values for each sensor. For example, green dots mean that the sensor is in the "normal" range. Red, on the other hand, represents one of the extreme categories. This permits a compact way of displaying a large amount of data and, if the dots are arranged to match the location of the sensor modules, they will show the pattern of the "exceptions". More detailed displays can be used to separate different types of sensors. Since the data are stored, it will be possible to review the past history and analyze the temporal pattern of the exceptions, as well as their spatial organization.

This approach using exception reporting is expected to help the sensor modules conserve power. It also forces early attention to examining preliminary data, finding appropriate theory, and building models to be used for the categorization. All too often these tasks are left to end, after the data have been collected. In the exception-reporting scheme, data analysis is done early and monitored closely.

Data summarization techniques are available for categorized data, such as consolidating the data with Theissen polygons (also called Delaunay triangulation) to show the general pattern of the data values as a map. Software such as GMT is available to produce this type of display. GMT, the Generic Mapping tools, is software written by Paul Wessel and Walter H. F. Smith, and is available at http://gmt.soest.hawaii.edu.

The hourly high-resolution images have a resolution of 1600 x 1200 pixels and serve several important interpretive functions. Viewed on a large-screen monitor, such as a wall-mounted, flat-plasma display, they provide a view such as you would get out a window onto the study site. This permits casual observations during periods where environmental conditions are reported as normal. During exceptional periods, this "window" provides an important visual check on the conditions and permits a quick analysis of how the vegetation is responding.

The images also provide a primary record of the life-history activities. Most of the images will be taken close to the rare or threatened plant species. This will permit observations of flowering, fruit set, fruit disappearance, leaf flushes, leaf loss and other significant events. Since the images are stored, it is easy to review them to confirm observations or review periods that were not being monitored. This type of data is usually very difficult to obtain by conventional visits to make observations.

7 Summary and Conclusions

We have been building a wireless network of environmental sensors used to monitor ecological conditions near endangered plants. We gather information at many points in the environment of the endangered plants and we also sample the information in surrounding environments where such plants do not grow. We focus on collecting light, wind, and a variety of rain and fog information. We also take high-resolution digital pictures of selected plants and their environment to monitor changes that our other sensors do not reveal.

The sensors and communication units must be reliable and consume very little power. Some of this is based on conventional techniques, such as having the units "sleep" as much as possible and the use of the highest energy density available in commercial batteries. Our novel techniques include the use of specialized routing algorithms, specialized sensors for detecting moisture, and special enclosures that will be overlooked by humans and will not harm the environment.

Given a large network of sensors continuously providing data, we find that one of our challenges is analyzing the large volume of data and providing ways for humans to derive useful information from this data. Our approach includes automatically generating thematic maps reflecting the relationship of the current situation to historical data, and computing models of the environmental conditions that allow us to predict expected future values and also flag for human attention values that do not match the model.

ACKNOWLEDGMENTS

This research is supported by a DARPA SensIT grant to the authors and Brian Chee (University of Hawaii at Manoa, Department of Information and Computer Sciences). We would like to acknowledge the design and construction skills of Michael Lurvey, the creator of our "not obvious" containers. Shu Chen is responsible for the design of MOR, and Wei Chen has collaborated with one of the authors on the design of Geometric Routing. Brian Chee provided essential guidance on many fronts and has kept our team of ICS undergraduate assistants well motivated. We would like to thank them, along with all of the students who learned how to make rocks in our Botany lab. Keith Bromley made a critical suggestion at the start of our research by urging us to seek ways to collaborate with other researchers. He was right and we have benefited greatly from working with Andy Wheeler and Rob Poor from Dr. Michael Hawley's Personal Information Architecture Laboratory at the MIT Media Lab. We also recognize the huge effort that other students from that lab made to get field experiments underway. Chris Newell's organizational skills kept everyone nearly sane during periods with critical deadlines.

BIOGRAPHIES

Edo Biagioni, Ph.D. is an Assistant Professor in the Information and Computer Sciences Department at the University of Hawaii at Manoa. He teaches undergraduate and graduate courses in computer networks, and supervises research projects and theses in protocol design and development, operating systems design, network training, Internet size and distance learning. He was instrumental in the development of the ATM protocol while working with FORE Systems, Inc. His advanced research into the develop of protocol stacks in high level languages while at Carnegie Mellon University provides an important foundation for research into problems involved with the use of communication protocols to help optimize energy budgets.

Kim Bridges, Ph.D. is an Associate Professor in the Botany Department of the University of Hawai`i at Manoa. He teaches courses in field ecology, as well as introductory general biology courses. He has also taught the introductory computer science course in the ICS Department. His background includes more than thirty years of experience with computer graphics and programming. He has previously taught courses in computer cartography and statistics. Dr. Bridges has published two books on the application of computer graphics to problems of data analysis and is an editor of a book on the ecology of Hawai`i Volcanoes National Park. His current research involves applying many facilities of the World Wide Web to problems of education and conservation research. He is a recent recipient of the Regents Medal for Teaching Excellence.

REFERENCES

Broch, J., Maltz, D. A., Johnson, D. B., Hu, Y. C., and Jetcheva, J. 1998. A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proceedings of the ACM/IEEE MobiCom*.

Feeney, L. and Nilsson, M. 2001. Investigating the energy consumption of a wireless network interface in an ad hoc networking environment. In *Proceedings of IEEE INFOCOM*.

Imielinski, T. and Navas, J. 1996. *GPS-based Addressing and Routing*. RFC 2009.

Johnson, D. B. and Maltz, D.A. 1996. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*, ed. T. Imielinski and H. Korth. Dordrecht: Kluwer, ch 5, pp. 153–181.

Jones, C. E., Sivalingam, K. M., Agrawal, P., and Chen, J. 2001, A survey of energy efficient network protocols for wireless and mobile networks. *ACM Journal on Wireless Networks* 7: 343–358.

Nievergelt, J. and Hinrichs, K. 1993. *Algorithms and Data Structures*, Englewood Cliffs. NJ: Prentice Hall, pp. 311–313.

Perkins, C. and Bhagwat, P. 1994. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. In *Proceedings of SigCOMM '94, London*. ACM, pp. 234–244.

U.S. Endangered Species Act of 1973 (PL 93-205; 16 USC 1531 et seq., as amended).