

# FloodNet: Coupling Adaptive Sampling with Energy Aware Routing in a Flood Warning System

Jing Zhou<sup>1</sup> (周 菁) and David De Roure<sup>2</sup>

<sup>1</sup>*School of Computer and Software, Communication University of China, Beijing 100024, China*

<sup>2</sup>*School of Electronics and Computer Science, University of Southampton, U.K.*

E-mail: zhoujing@cuc.edu.cn; dder@ecs.soton.ac.uk

Received January 10, 2006; revised September 28, 2006.

**Abstract** We describe the design of FloodNet, a flood warning system, which uses a grid-based flood predictor model developed by environmental experts to make flood predictions based on readings of water level collected by a set of sensor nodes. To optimize battery consumption, the reporting frequency of sensor nodes is required to be adaptive to local conditions as well as the flood predictor model. We therefore propose an energy aware routing protocol which allows sensor nodes to consume energy according to this need. This system is notable both for the adaptive sampling regime and the methodology adopted in the design of the adaptive behavior, which involved development of simulation tools and very close collaboration with environmental experts.

**Keywords** adaptive sampling, communication system routing, energy awareness, environmental sensor networks

## 1 Introduction

Pervasive computing devices provide exciting new opportunities for intelligent sensing and monitoring the natural environment, such as measuring water levels and air pollution. Significantly, these technologies make it possible to deploy more devices in order to obtain more data more often, and this greater richness of data is set to create a powerful impact on environmental monitoring and decision-making. Traditional solutions involve data loggers from which data is collected periodically in person or via telemetry. With wireless communications and energy drawn from local sources such as solar cells, devices can be deployed without the constraints of having to wire them up or make them accessible, and data can be conveyed when needed.

Achieving these benefits within the natural environment brings a number of technological challenges. It is often also the case that deployment is expensive and it may be difficult or impossible to access the devices again later, in contrast to working with handheld pervasive devices. Devices need to withstand harsh environmental conditions. There may also be large numbers of nodes and they need to have coordinated behavior, but at the same time we have to assume that node failure, temporary or permanent, is likely to occur — the natural environment is a place of change.

One of the most substantial challenges, and our focus here, is optimizing the consumption of electrical power. At times when stored power is not being fully replenished, the more often data is sent then more likely it is that the device will not have sufficient power to continue its function.

In this paper we describe a flood early warning system, FloodNet. Flood damage represents a major on-

going cost and risk may be increasing due to land-use

change, climate change and flood-prone investment. When a flood occurs, the cost of damage has a clear correlation with both the depth of the flooding and the time in advance at which warning is given. By applying pervasive computing technologies on the floodplain we have the potential to obtain better data from which to make predictions, and we can do this in a timely manner in order to improve warning times. Such systems provide an excellent illustration of the benefits and challenges of pervasive computing in the environment. Deployment is facilitated by wireless technologies but we have issues of power for the devices and the need for very long unattended periods of operation. This scenario also emphasizes the energy optimization challenge, because the data is most important during flood conditions and this is when solar energy is typically least available.

The fundamental tradeoff between the need for timely data and the need to conserve energy is the research focus of the project. The goal is to make the system adaptive so that the sampling and reporting rates of the devices vary according to need, conserving power and minimizing the data volume required. Some intelligence is required in this adaptation, because the importance of a device at a given moment will depend upon both its environmental context and its role in relaying data from other devices, each of which will vary dynamically according to circumstances. The key element of our strategy is the use of a centralized flood predictor model (see Section 2) which was developed by environmental experts. From this the priorities for collecting samples from each sensor can be determined. Reducing activity of individual sensor nodes and minimizing the data volume required help prolong network

lifetime; i.e., the time until the network fails<sup>①</sup>. The use of the predictor model within this adaptive sampling regime is a distinctive feature of this work. This new technology in the field makes new demands on the back-end processing. The higher temporal and spatial density of data, and faster turnaround times, present a significantly greater computational task in the flood modelling. For this reason we turn to Grid computing for the back-end processing. This combination of Grid and Pervasive computing is another distinctive feature.

The system is illustrated in Fig.1, which also shows the possible control loops in the adaptive sampling regime. The outermost loop enables the flood predictions to influence the reporting rates of individual nodes, so that closer monitoring can be achieved in anticipation of a possible flooding event. The inner loop is the peer-to-peer behavior of a set of nodes which can communicate with each other but have no external coordination and are therefore described as *self-managing* or *autonomic* (a word borrowed from the notion of the autonomic nervous system). Other possibilities include one node, such as the gateway, taking a coordinating role.

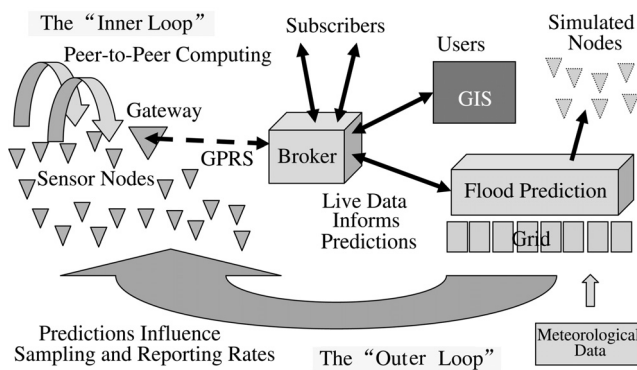


Fig.1. Adaptive sampling control loops.

Obtaining the required behavior from a system of this complexity is a challenge. Our third distinctive is our methodology for designing the adaptive sampling behaviors in this hybrid Grid-Pervasive system. We worked closely with environmental scientists in order to understand the data prioritization — this is non-trivial, because the significance of the data reported by a node is a function of its location and varies depending on the conditions reported by other nodes locally and from other sites. Rather than adopting established energy-conservation protocols, the adaptive sampling behavior in our system is the result of an extensive co-design exercise with environmental experts.

In the next section we outline the FloodNet project. We then describe the energy aware, adaptive routing in FloodNet in Section 3 and in Section 4 we present the evaluation methodology and results. The review of related work on adaptive sampling and energy aware

routing is given in Section 5. We close by reflecting on the lessons learned in Section 6.

## 2 Project Background

The FloodNet project has deployed a set of intelligent sensor nodes around a stretch of the River Crouch (Essex, UK). This site was chosen for its tidal behavior so that, for test purposes, there are regular variations in water level. The nodes are powered by solar cells in conjunction with batteries and each node communicates with its neighbors using wireless Ethernet. A special node, the gateway, relays the data back to base using GPRS (General Packet Radio Service). This is an ad hoc network, with nodes relaying information across the network to ensure data delivery to the gateway. Various parties can subscribe to the incoming data stream. As well as being stored in a GIS (Geographical Information System), the data is used to inform flood simulations which are used to make flood predictions. The topology of the sensor network is depicted in Fig.2. One can envisage a number of such deployments at different locations in the river, each reporting back in this way — the current deployment enables us to explore the issues of working with this spatial density of data.

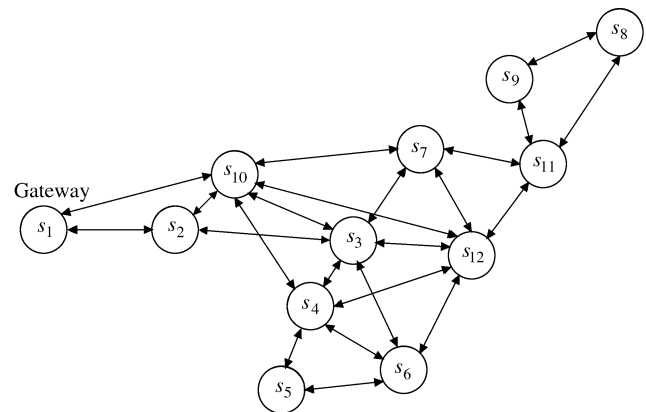


Fig.2. FloodNet sensor network.

The process of adaptive sampling is mediated by the use of a flood predictor model<sup>[1]</sup> which comprises a stochastic one-dimensional numerical hydraulic model (ISIS) coupled to an ensemble Kalman filter. The model allows for the real-time collection of water depth data to update the flood predictions regularly with refreshed data. When the model-based probability of the water level exceeding a threshold at a validation location is less than 5%, the requirement for transmission of data from the sensor node at the validation location is lowered. Otherwise, the requirement for transmission of data will arise. The degree of the model requirement is represented by means of *data importance*. The predictor model is required to carry out extensive processing in a short period of time (currently 1 hour) to produce

<sup>①</sup>In the context of FloodNet, the network failure means that data messages cannot be generated by sensor nodes or cannot be retransmitted due to energy depletion.

the data importance for sensor nodes during the following iteration. Upon each model iteration the network changes its behavior, altering the reporting rate (derived from the data importance using a conversion function specified by the environmental experts) of each individual node according to the data importance placed upon it by the predictor model (see Fig.3). In the broadest sense the nodes that first experience a flood event will have a high initial demand. This demand will ease as the model develops a sufficient level of confidence in the prediction for this particular node and as the flood event itself penetrates into surrounding areas, raising the data demand from these areas. With time the more disparate nodes become more active. As such a wave of activity passes from the sensor nodes in the main channel out to the floodplain areas.

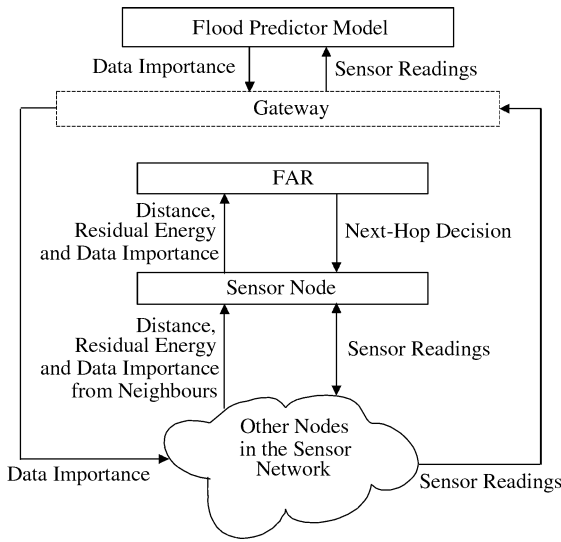


Fig.3. Information flow.

### 3 FloodNet Adaptive Routing

To extend the lifetime of wireless sensor networks, many routing protocols have been proposed with energy awareness being an essential design consideration. These schemes involve selecting the optimal path between a data originator and a data collecting unit. The selection criteria take into account the available power in the nodes or the energy required for transmission in the links along the path<sup>[2]</sup>. Some other conventional optimality criteria include hop count, delay, forwarding cost on the link, etc. Unlike others (see Section 5), in FloodNet the flood predictor model influences the node reporting frequency. We take this feature into consideration during protocol design. The resulting energy aware routing for FloodNet has more complex optimality criteria for path selection, coupling application requirements with energy conditions in the network.

Here we provide an overview of the FloodNet adaptive routing (FAR), consisting of interest diffusion, neighbor status maintenance and the adaptive routing

algorithm. In our study, nodes are not mobile and the network topology is static. We assume constant transmission power for sensor nodes.

#### 3.1 Interest Diffusion

The predictor model requirement (aka interest), is diffused through the network from the gateway using a flooding technique on an hourly basis (see Fig.2). This process helps establish *neighbor tables* at individual nodes, which capture neighbor information for these nodes.

The description of an interest consists of a list of data importance placed upon each node the value of which is within range  $[0, 1]$ . An interest message comprises the following fields: the unique identifier and the residual battery power of the sender, the message type, the message timestamp, the distance of the message recipient from the gateway, and the data importance for each node. The data importance will be converted locally by its associated sensor node to one of the three reporting rates (currently once, twice and three times per model iteration) allowed.

If an incoming message is regarding a fresh interest, then a sensor node keeps a copy in its *interest cache*. For messages regarding the same interest but from distinct senders, the node extracts related information to fill in its neighbor table (see Table 1), replaces the sender identifier and residual battery power in the message with its own information, and increases the distance in the message by 1. Then, the node only broadcasts the first message of an interest to others within its wireless transmission range. This process is carried out recursively to ensure that each node will be notified of such an interest. Table 1 illustrates the neighbor table of node  $s_7$  in which the information about its neighbors has been filled in. The residual battery power of neighbors will be regularly updated through neighbor status maintenance as presented in the following section. The link cost is assumed to be known to the involved pair of sensor nodes.

Table 1. Neighbor Table of Node  $s_7$

Node Identifier	Distance	Residual Battery	Data Importance	Link Cost
$s_{10}$	1	0.91	0.70	$e_{7,10}$
$s_3$	2	0.93	0.85	$e_{7,3}$
$s_{12}$	2	0.95	0.90	$e_{7,12}$
$s_{11}$	3	0.98	0.90	$e_{7,11}$

#### 3.2 Neighbor Status Maintenance

The neighbor status comprises the unique identifier, the distance from the gateway, the residual battery power and the data importance of a neighbor. Sensor nodes maintain this information in neighbor tables to assist data message forwarding. As mentioned earlier, the neighbor table is set up and populated by extracting the neighbor information from the interest message at the start of each model iteration. During each model iteration, among others, only the residual battery power

may be changing and hence information about it needs to be refreshed to remain up-to-date.

The frequency on which a sensor node broadcasts its up-to-date information is a protocol design parameter that is determined by the accuracy requirement of the component that will utilize the information. As presented later in Subsection 3.3 the adaptive routing algorithm will take into account the neighbor status to decide the optimal path for routing, the frequency of local broadcast should therefore correspond with the reporting rate of nodes because routing decision needs to be made once reporting activities commence.

The three reporting rates allowed for FloodNet sensor nodes indicate that reporting could occur at the 20th, 30th, 40th, and 60th minute of each model iteration (lasts for 1 hour), and local broadcast of neighbor status is triggered prior to the start of any reporting activity<sup>②</sup>. Sensor nodes receive updates of neighbor status and modify their neighbor tables wherever necessary.

### 3.3 Adaptive Routing Algorithm

The adaptive routing algorithm we proposed is used for making a decision on which neighbor a sensor node should forward a data message to. We use a metric, *priority*, to denote the degree that a sensor node should be chosen to forward the data message. Ideally, the greater the priority of a node is, the more likely the node would be selected. The priority of sensor node  $s_1$  with respect to  $s_0$  has the following form:

$$p_0^1 = \frac{b_1}{e_{0,1} \times \beta^{t_1}}$$

where  $b_1$  represents the residual battery power of  $s_1$ ,  $e_{0,1}$  is the link cost, i.e., the energy required for transmission on link  $s_0 \leftrightarrow s_1$ ,  $t_1$  denotes the data importance of  $s_1$ , and  $\beta$  ( $\geq 1$ ) is a tunable parameter<sup>③</sup>.

In the cases where a few sensor nodes share the same priority the following rules should be applied:

**If** nodes have the same priority and different data importance, **then**  $s_0$  shall choose the one with the least importance;  
**else if** nodes have the same priority and the same data importance, **then**  $s_0$  shall choose the one with the least distance;  
**else if** nodes have the same priority, data importance and distance, **then**  $s_0$  shall randomly choose one.

We present the details of the adaptive routing algorithm as follows.

1) A sensor node sends a data message to a neighbor with the highest priority among all neighbors closer to the destination.

2) If no such neighbors exist, the node sends the data message to a neighbor with the highest priority among

all neighbors that are of the same distance from the destination.

3) If all neighbors are farther from the destination, the node sends the data message to a neighbor with the highest priority among all neighbors; otherwise, the data message is dropped.

4) Steps 1, 2 and 3 are repeated until the data message reaches the gateway, or is dropped out.

Each data message carries the following information: the sender and the receiver identifiers, the message type, the message timestamp, the sensor readings and a list of visited nodes. Hence, forwarding loops are prevented as a sender will always check the list of visited nodes in the data message before retransmitting it.

The priority function has been carefully chosen to conserve the battery power of nodes with more important data. It produces a higher priority, which means more chances of being used as a router for data transmission, for sensor nodes with ample battery power and a light reporting task, whilst giving a lower priority to those with a lower level of battery power and a heavy reporting task. Furthermore, the function can prevent the battery power of some nodes from being drained much earlier than that of others, thus delaying the occurrence of energy depletion which leads to data loss. Suppose  $s_0$  has two neighbors  $s_1$  and  $s_2$  with  $t_1 < t_2$ . Let  $c_1 = \frac{b_1}{e_{0,1}}$ . If  $c_2$  is far greater than  $c_1$ , i.e.,  $c_2 - c_1 > c_{threshold}$ , it is reasonable to choose  $s_2$  as the next-hop node especially when  $s_1$  has very little residual energy or  $s_1$  is associated with a very high link cost. With the priority function,  $s_2$  will be selected as the next-hop node if the following inequality holds:

$$c_2 > c_1 \times \beta^{t_2 - t_1} \geq c_1.$$

In the FloodNet adaptive routing, delivery of data messages has no dependency on any specific node, as each sensor node maintains a neighbor table listing pointers to multiple neighbors which can further relay data back to the destination, and the optimal path is computed on demand. Hence, the adaptive routing protocol is robust to topological changes due to transient node and link failure.

## 4 Performance Evaluation

### 4.1 Use of Simulators

The fundamental purpose of the sensor network is to provide information to the appropriate quality demanded by the users. Hence in designing the system we focused on the requirements of the users of the information, and in particular with those familiar with the site. Through observing discussions with the users

<sup>②</sup>During interest diffusion, neighbor status has been updated through interest messages. Hence, local broadcast of neighbor status within the model iteration would only need to occur prior to the 30th, 40th and 60th minute.

<sup>③</sup> The discussion about the best value range for  $\beta$  can be referred to in Subsection 4.2.

we noted that they were expressing a considerable degree of knowledge about the environmental situation of each node, about priorities and redundancy of data, and about desired behaviors in a range of different circumstances. It was evident that they sought a very sophisticated behavior from the system and brought considerable knowledge. To capture this knowledge and to evaluate the designed system, we constructed two simulation tools.

#### 4.1.1 High-Level Simulator

The first was a high-level tool which allowed us to express behaviors at a similar level to the statements that the users were making — essentially capturing rules. This enabled us to establish the essential features of the adaptive routing protocol (see Section 3). Although we experimented with an expert-system approach (using JESS, the Java Expert System Shell), we found it useful to build a custom design tool in Scheme<sup>[3]</sup>. This was effectively an exercise in *metalinguistic abstraction*; i.e., we created a small language which was as close as possible to the ways in which we found the users were describing adaptive sampling behaviors.

The simulation approach was based on [4], with a small number of primitive operations to capture the essence of the sensor network and thereby raise the relevant issues. A node was defined to have a number of local state variables and to be able to communicate through a simple broadcast to neighbors within a specified radius and with a specified probability of success. The nodes were required to bootstrap themselves from an initial situation in which they had no knowledge of their neighbors. Behavior was expressed as a set of rules which are triggered on incoming messages according to local state. The basic inter-node messages were:

*HELO*: a null message for anyone who is listening so that they know they are receiving from this node identifier;

*HOP*: announces to neighbors the number of hops from this node to the gateway;

*STATUS*: announces to neighbors the node status, with parameters including battery level and data importance;

*LEVEL*: relayed message which carries sensor readings.

Some of the “intelligence” was encoded in a *route* function which describes the next node to which an incoming message is to be relayed based on a local neighbor table which is updated according to incoming messages. Constructing this simulator and performing experiments with it give insights into engineering a more comprehensive discrete-event simulator described in the following section.

#### 4.1.2 Discrete-Event Simulator

A second simulator was used to conduct a full simulation of the deployed network. The simulator was implemented in Java and deployed on a single machine. Each

sensor node was effected as an object and communicated with each other through message passing mechanisms. The simulator provided an input message queue to host the incoming messages of all nodes to be processed and a simulator clock to capture the advance of simulation time at each node. All the messages were stored and processed in non-decreasing order of their timestamps.

FAR is an upper layer (network) protocol that assumes a collision-free, multiple retry MAC layer. The simulator implemented the basic FAR protocol. In addition to a number of parameter choices in the routing protocol, we can also vary certain factors in simulation such as the number of sensor nodes, the pattern of power consumption of sensor nodes, and the frequency of reporting activities of sensor nodes. The simulator enabled us to explore various “what-if scenarios”, and to address questions of scalability. It also validated the algorithm ready for deployment.

## 4.2 Experimental Settings

The FAR performance was evaluated through simulation, using the second simulator described in Subsection 4.1.2. The basic simulation parameters were chosen to model an ad hoc network which consists of 12 sensor nodes with 1 of them residing on the gateway. The initial battery power for all nodes is 1 unit. The exception to this is  $s_1$  on the gateway which was assumed to have ample energy at all times. We assumed constant transmission power for sensor nodes. The energy consumption during idle time was not included in the design and the simulation as we assume a situation under which the energy consumption due to interest diffusion, neighbor status maintenance and data message delivery is dominant. The battery consumptions of full power transmitting and receiving are 0.005 unit and 0.0005 unit. The data importance for each node is randomly distributed over  $[0, 1]$ . Each model iteration is simulated to last for 1 hour. Having different reporting rates, sensor nodes may generate and send data messages at the 20th, 30th, 40th and 60th minute of each model iteration which correspond to the 4 *time units* of one simulated hour.

The simulator implemented another protocol we refer to as the minimum energy consumption forwarding (MECF) for comparison purposes. The distinction between FAR and MECF is that MECF only forwards data messages to a neighbor which leads to a minimum energy consumption path, i.e., the priority in MECF is calculated on the basis of the distance of neighbors and the associated link cost. We anticipate to see the improvements that FAR brings to its performance by taking into account the residual battery power and the data importance of sensor nodes.

Moreover, the simulator implemented the Energy Aware Routing protocol<sup>[5]</sup> which aimed to increase the survivability of network. The weighting factors used in the experiments on EAR are  $\alpha = 1$  and  $\beta = 50$ , as used in [5].

**Table 2.** Comparison Between MECF and FAR with Different  $\beta$  in FloodNet Topology

	MECF	FAR (1.0)	FAR (1.0001)	FAR (1.001)	FAR (1.01)	FAR (1.1)	FAR (1.5)
NL	30.65	32.35	32.55	32.00	32.35	31.90	31.55
NOT	30.20	31.95	32.25	31.85	31.95	31.60	31.40

**Table 3.** Comparison Between MECF and FAR with Different  $\beta$  in Random Topologies

	MECF	FAR (1.0)	FAR (1.0005)	FAR (1.005)	FAR (1.05)	FAR (1.5)
NL	38.95	41.10	41.70	41.30	41.20	40.85
NOT	38.85	41.05	41.35	40.95	40.90	40.40

**Table 4.** Network Lifetime and Node Operational Time for FAR ( $\beta = 1.0005$ ) and EAR

# Nodes	12				24			
Topology	FloodNet		Random		FloodNet-Like		Random	
Protocol	EAR	FAR	EAR	FAR	EAR	FAR	EAR	FAR
NL	29.75	32.55	39.20	41.70	15.85	16.20	26.65	30.30
NOT	29.45	32.25	38.85	41.35	15.80	16.20	26.05	29.80

To study the performance increase of FAR over EAR as a function of network density, we also generated sensor fields (30 units by 30 units) in which different numbers of nodes are randomly placed. The number of sensor nodes ranges from 8 to 24 nodes in increments of 4 nodes and each node has a transmission range of 15 units.

We measured the *success ratio* which is defined as the ratio between the number of data messages successfully received by the gateway to the total number of data messages that should be generated by all sensor nodes in the field. We also measured the *network lifetime* which we define as the first time unit at which data messages cannot be generated by sensor nodes or get lost due to insufficient energy. The *node operational time* is defined as the first time unit at which energy of any of the sensor nodes is depleted. For exploring the robustness of FAR, link failure was simulated by a probability of failing links that occur randomly in each model iteration. We executed 20 runs of the simulator for each of different protocols and of a number of different tunable weight  $\beta$  (in FAR) in the simulated ad hoc network<sup>④</sup>.

### 4.3 Results

The comparison between MECF and FAR with different  $\beta$  in the FloodNet topology is shown in Table 2. It can be observed that FAR outperforms MECF in all cases by producing a longer network lifetime (NL) and a longer node operational time (NOT). We believe, by carefully choosing a neighbor to forward data messages, FAR enables the energy of sensor nodes to be consumed according to need whilst striking a balance between minimizing the overall energy consumption and maximizing the minimum residual battery power in sensor nodes. Hence, it extends NL and NOT.

Also, we carried out experiments on MECF and FAR in a number of random topologies ( $\geq 20$ ) and Table 3 re-

veals the result. Both MECF and FAR deliver a better performance in random topologies than in the FloodNet topology. FAR achieves its best performance in the FloodNet topology with  $\beta = 1.0001$  and in the random topologies with  $\beta = 1.0005$ , which demonstrates that such small values for  $\beta$  can prevent the energy of some nodes being depleted far earlier than that of others. The best value for  $\beta$ , however, may need to be tuned to the particular example scenario.

Table 4<sup>⑤</sup> presents the network lifetime and the node operational time that FAR ( $\beta = 1.0005$ ) and EAR can achieve respectively. FAR outperforms EAR in all cases that differ in the number of nodes in the network and/or the network topology. In the FloodNet topology with 12 nodes, FAR extends NL by 9.4% and NOT by 9.5% over EAR. Both protocols deliver a longer NL and a longer NOT in the random topologies than in the FloodNet topology. The reason for this is that in random topologies the outgoing links are more evenly distributed among sensor nodes than in the FloodNet topology and consequently, the probability that the energy of the nodes closer to the gateway is depleted much earlier than that of others, is reduced. As the number of sensor nodes increases, FAR produces less increase in NL and NOT over EAR. The corresponding result on the success ratio is depicted in Fig.4. FAR outperforms EAR by delivering a higher success ratio in all the topologies involved. Furthermore, as the number of the sensor nodes increases, both NL and NOT are reduced (see Table 4), thus leading to a lower success ratio.

Fig.5 plots the success ratio as a function of the link failure rate in an ad hoc network of different topologies. FAR delivers a higher success ratio in the random topologies than in the FloodNet topology. In the FloodNet topology, the success ratio is above 90% for the link failure rate up to 28%. In a small-sized ad hoc network like FloodNet, the functioning of certain links such as  $s_0 \leftrightarrow s_2$  and  $s_0 \leftrightarrow s_{10}$ , is more crucial to a high success

<sup>④</sup>It should be noted that being situated on the gateway,  $s_1$  is treated differently from others and all simulation results presented in this section do not include data of  $s_1$ .

<sup>⑤</sup>The FloodNet-like topology with 24 nodes corresponds to a combination of two 12-node FloodNet topologies as in Fig.2.

ratio than that of others, as no data messages can be successfully delivered to the gateway if they both fail. In the random topologies, when 39% links fail, the success ratio remains above 90%. We believe the robustness of FAR is attributed to its routing algorithm as the latter always tries to find alternative forwarding neighbors even if there is a hole (i.e., no nearer neighbor exists) that could be caused by the failure of either links or nodes.

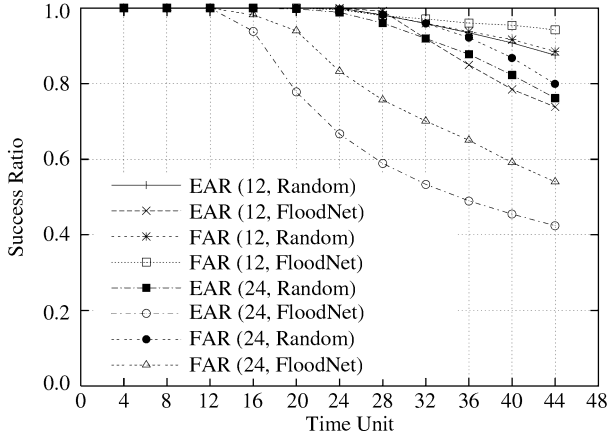


Fig.4. Success ratio for FAR ( $\beta = 1.0005$ ) and EAR.

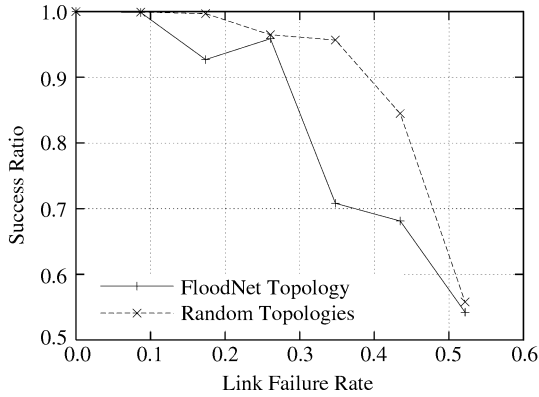


Fig.5. Success ratio for FAR ( $\beta = 1.0005$ ) with different link failure rates.

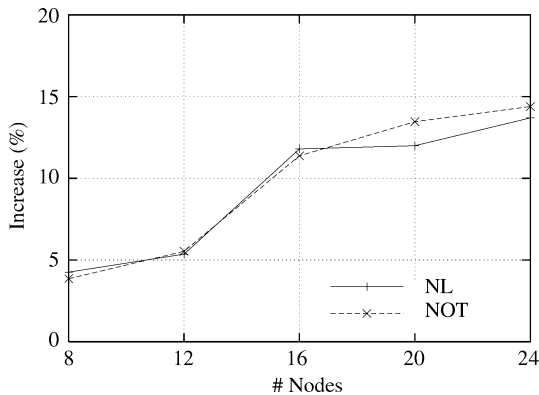


Fig.6. Effect of network density on performance increase of FAR over EAR in random topologies.

Fig.6 shows the effect of network density on the performance increase of FAR over EAR in random topologies. FAR achieves a little increase (around 4%) in both NL and NOT in a low density network (with 8 nodes). As the network density becomes higher, FAR will deliver a better performance. An increased network density means more neighbors per node on average. Hence, FAR will have more candidate nodes to choose from to determine a desirable routing path. By examining the simulation traces we found that the improved performance by FAR was obtained at the cost of more overall energy consumption. This is because the routing decision may have to involve a longer path in order to extend network lifetime whilst satisfying the predictor model requirement.

#### 4.4 Discussion

We employed a flooding technique in FAR to propagate the predictor model requirement in the sensor network, which resembles interest diffusion in directed diffusion<sup>[6]</sup>. In directed diffusion, the sink is required to periodically refresh the interest to increase delivery reliability. We acknowledge that the successful delivery of interest messages in our scenario is very important and there are existing methods (e.g., [7]) which guarantee the delivery of messages and might be suitable for FloodNet. However, as interest diffusion is mainly intended to find the distance information for each node, applying those techniques for guaranteed delivery may require non-trivial modifications to accommodate this requirement. We found through simulation that flooding is sufficient for interest diffusion in FloodNet of its current scale and form. Definitely, we will need to investigate the issue of guaranteed delivery as FloodNet evolves. Note that due to the inherent nonscalability of the flooding technique and the neighbor table driven approach, the proposed solution may not be scalable enough for direct application to large scale sensor networks.

As a reactive routing protocol, FAR computes routes on demand and only needs to maintain routes to sensor nodes required. It was mentioned earlier that sensor nodes in the FloodNet scenario were deployed at a stretch of river, and thus (moving) obstacles in the signal path may lead to the loss of communication links. Also, some nodes may be temporarily inactive. We believe it is better to use a reactive protocol in FloodNet since proactive routing protocols require the knowledge of node activity to make decisions.

FAR relies on the concept of localized routing algorithms<sup>[8]</sup> in which sensor nodes only communicate with nodes within some neighborhood. Routing decisions are made on the basis of the priority derived from local information and the model requirement. Achieving a desirable global objective is essential for the success of such algorithms. Unlike many other energy aware protocols, FAR takes into consideration the requirement

that the reporting frequency of nodes should be adaptive to both local conditions and the predictor model as needed by FloodNet. Hence, the criteria for determining the best routing path should not depend upon the energy consumption on the path only<sup>⑥</sup>. This also explains why our definition for the network lifetime differs from that of others, for example [10] defined it as the time until the network partition occurs due to battery depletion whilst [11] modelled it as the earliest time that a message cannot be sent.

Singh *et al.*<sup>[9]</sup> argued that it may be a better solution to route packets along nodes having sufficient energy, or nodes under light loads, to increase node and network life. This implies that an optimal path in terms of energy awareness is not necessarily the shortest-hop path. As mentioned in Subsection 4.3 we did observe that routing in FloodNet occasionally takes a longer route, which potentially increases the energy consumption. Essentially, our approach trades off the overall energy consumption for energy aware, adaptive behaviors of sensor nodes. This is why although link cost between sensor nodes is considered in the priority function, it is not as significant as other factors, such as the residual energy, the data importance, and the distance, in contributing to the selection of the next-hop node for data delivery.

## 5 Related Work

### 5.1 Adaptive Sampling

In view of the limited resources in wireless sensor networks, adaptive sampling has been introduced to handle this issue by making the data collection rate dynamic and adaptable to the environment. We briefly review some of the related work in this section.

Willett *et al.*<sup>[12]</sup> proposed a backcasting approach to reduce energy consumption whilst maintaining highly accurate measurements of sensors. Initially, a subset of the sensor nodes communicate their information to a fusion center. This provides an estimate of the environment being sensed, indicating some sensors may not be required to achieve a desired level of accuracy. The fusion center then backcasts information based on the estimate to the network at large and selectively activates additional sensors to obtain a target error level. In this approach, adaptive sampling can save energy by only activating a fraction of the available sensors.

Marbini and Sacks presented in [13] a feedback control mechanism which makes the rate at which each sensor collects readings from the environment dynamic and adaptable. The sampled data are compared against a model that represents the environment. An error value is calculated based on this comparison. If the error value is more than the predefined margin of error, then the sensor node will collect data at a higher sampling rate; otherwise, the sampling rate will be decreased. Sensor

nodes are completely autonomous in adapting the sampling rate.

Jain and Chang<sup>[14]</sup> developed an adaptive sampling scheme which adjusts data collection rate in response to the characteristics of the streaming data. A Kalman filter is used at each sensor to make predictions of future values of a data stream based on those seen. The sampling interval (SI) is adjusted based on the prediction error. If the modification in SI exceeds that allowed by the specified Sampling Interval Range (SIR), a new SI is requested from the server. At the central server, new SIs are allocated to requesting sensors according to available bandwidth, network congestion and streaming source priority. Sensors that monitor objects showing increased activity will have low SI whilst those with reduced activity will have high SI.

### 5.2 Energy Aware Routing

In this subsection, we describe and compare existing energy aware routing protocols with FAR. To the best of our knowledge, none of the related work described here examines the impact that the distinct reporting rates have on the protocol design.

Stojmenović and Lin<sup>[15]</sup> proposed power-aware localized routing which only requires localized routing information to minimize the energy consumption and extend battery's worse case lifetime. However, the performance of their method was demonstrated in networks with high connectivity. Further, the FloodNet sensor nodes are not equipped with any mechanisms to provide location information, the GPS (Global Positioning System) receivers for example. Hence we are unable to directly apply their approach to FloodNet. Other energy aware routing protocols using the location information of nodes include GEAR<sup>[16]</sup> which achieves longer network lifetime by balancing energy usage across the network. GEAR uses energy aware and geographically informed neighbor selection heuristics to forward a data packet towards the target region and applies recursive geographic forwarding or restricted flooding to disseminate the data packet within the region.

To maximize the network lifetime, Chang and Tassiulas<sup>[10]</sup> proposed formulating the routing problem in wireless sensor networks consisting of static nodes as a linear programming problem for constant data rates as well as arbitrary processes. The major assumption in their work is that global network information is available for decision making on routing. This assumption is also shared by [11]. Hence, each origin node of a commodity can calculate the shortest cost path to its destination node. This is, however, not suitable for the FloodNet scenario in which each node only has a local knowledge of the global network.

LEACH<sup>[17]</sup>, PEGASIS<sup>[18]</sup> and the Energy Aware Routing protocol<sup>[5]</sup> extend network lifetime by evenly distributing the energy load among all the nodes in the

<sup>⑥</sup>A set of different power aware metrics to determine the optimal paths in mobile ad hoc networks were presented in [9].



network. They assume that nodes have a fixed sampling and reporting rate. Moreover, LEACH and PEGASIS assume that nodes are able to transmit with enough battery power to reach the base station. Although there are superficial similarities between LEACH and FAR, and specifically LEACH also considers non-uniform energy consumption of nodes, e.g., cluster-heads often consume more energy for data forwarding, LEACH essentially achieves a longer network lifetime by evenly dissipating energy among all nodes whilst FAR does this by enabling nodes to consume energy according to need (see Section 1).

Schurgers *et al.*<sup>[19]</sup> proposed using topology management techniques to coordinate the sleep transitions of all nodes whilst ensuring adequate network connectivity. The sensor network was assumed to be in the monitoring state during most of its lifetime. Power savings were achieved at the cost of an increased path setup latency. This scenario also motivated [20], an energy-efficient MAC protocol which tries to reduce the waste of energy from collision, overhearing, control package overhead and idle listening. In contrast, sensor nodes in FloodNet are required to report data at different but more frequent rates and thus these methods are not appropriate for solving the FloodNet problem.

The Pulse protocol<sup>[21]</sup> was designed for multi-hop wireless infrastructure access to mobile users which utilizes a periodic flood initiated at the gateways to provide the routing and synchronization information to the network. Substantial energy savings can be achieved by using the synchronization information to allow idle nodes to power off their radios for a majority of the time when they are not required for packet forwarding. The Pulse flood proactively maintains a route from all nodes in the network to the infrastructure access node. This is in contrast with FAR which selects the routing path on-the-fly as data messages traverse the network and thus is robust to temporary failure.

## 6 Conclusions and Future Work

FloodNet effectively demonstrates the potential of using pervasive computing in the environment and highlights solutions to important issues, such as adaptive sampling for prioritized data gathering and energy conservation. It has enabled innovation within the environmental application and has also provided an important case study to explore the relevant research and development issues in the computer science field. This project brings together two significant trends in the evolution of computing — pervasive computing and Grid computing — to build a sensor network application using a design methodology that might be called “extreme protocol design”. This has led to an interesting energy aware protocol with demonstrable benefits over other solutions.

At the time of writing, a second phase of nodes is about to be deployed. The completion of our experiment will be when we have collected sufficient environmental

data across varying conditions that we can confirm with the expert end-users that the adaptive sampling is performing as required. We will report this aspect of the work in a future paper.

Although nodes have well-defined local behavior (inner loop) and participate in the outer control loop, we have not explored node behaviors in which coordinated decisions are made across subsets of nodes which are isolated from the outer control loop; for example, through leader election. This is a topic which we are able to explore through the simulator.

We also need to conduct further work on isolation of nodes which have a disruptive behavior. One approach which we are able to explore through simulation is for a set of nodes to isolate a disruptive node by agreeing to communicate at different times.

**Acknowledgement** FloodNet is funded by the UK Department of Trade and Industry under the Next Wave Programme, as part of the EnviSense Centre for Pervasive Computing in the Environment. The project partners are University of Southampton, IBM, Multiple Access Communications Ltd, Halcrow Ltd and ABP Marine Environmental Research Ltd. We are particularly grateful to Craig Hutton of the Geodata Institute and to Jeff Neal and Peter Atkinson of the School of Geography for specifying the adaptive sampling requirements. We also acknowledge the input of the Equator Project (EPSRC Grant GR/N15986/01).

## References

- [1] Neal J, Atkinson P, Hutton C. Real-time flood modelling using spatially distributed dynamic depth sensors. In *Geosensornetze — von der Forschung zur praktischen Anwendung*, Brox C, Krüger A, Simonis I (eds.), Institut für Geoinformatik, Universität Münster, Germany, IfGI Prints, Vol.23, 2005, pp.147–159.
- [2] Akyildiz I F, Su W, Sankarasubramaniam Y, Cayirci E. A survey on sensor networks. *IEEE Communications Magazine*, August 2002, 40(8): 102–114.
- [3] Abelson H, Dybvig R K, Haynes C T *et al.* Revised report on the algorithmic language scheme. *Higher-Order and Symbolic Computation*, August 1998, 11(1): 7–105.
- [4] Adams S, De Roure D. A simulator for an amorphous computer. In *Proc. the 12th European Simulation Multiconference on Simulation — Past, Present and Future*, Manchester, United Kingdom, 1998, pp.153–160.
- [5] Shah R C, Rabaey J. Energy aware routing for low energy ad hoc sensor networks. In *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Orlando, FL, March 2002, pp.350–355.
- [6] Intanagonwiwat C, Govindan R, Estrin D. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *Proc. the 6th Annual Int. Conf. Mobile Computing and Networking*, Boston, Massachusetts, USA, 2000, pp.56–67.
- [7] Stojmenović I. Geocasting with guaranteed delivery in sensor networks. *IEEE Wireless Communications*, December 2004, 11(6): 29–37.
- [8] Estrin D, Govindan R, Heidemann J, Kumar S. Next century challenges: Scalable coordination in sensor networks. In *Proc. the 5th Annual ACM/IEEE Int. Conf. Mobile Computing and Networking*, Seattle, Washington, USA, 1999, pp.263–270.

- [9] Singh S, Woo M, Raghavendra C S. Power-aware routing in mobile ad hoc networks. In *Proc. the 4th Annual ACM/IEEE Int. Conf. Mobile Computing and Networking*, Dallas, Texas, USA, 1998, pp.181–190.
- [10] Chang J H, Tassiulas L. Maximum lifetime routing in wireless sensor networks. *IEEE/ACM Trans. Networking*, August 2004, 12(4): 609–619.
- [11] Li Q, Aslam J, Rus D. Online power-aware routing in wireless ad-hoc networks. In *Proc. the 7th Annual Int. Conf. Mobile Computing and Networking*, Rome, Italy, 2001, pp.97–107.
- [12] Willett R, Martin A, Nowak R. Backcasting: Adaptive sampling for sensor networks. In *Proc. the 3rd Int. Symp. Information Processing in Sensor Networks*, Poster Session, California, USA, 2004, ACM Press, pp.124–133.
- [13] Marbini A D, Sacks L E. Adaptive sampling mechanisms in sensor networks. In *Proc. London Communications Symposium 2003*, London, UK, September 2003. University College London.
- [14] Jain A, Chang E Y. Adaptive sampling for sensor networks. In *Proc. the 1st Int. Workshop on Data Management for Sensor Networks*, Toronto, Canada, 2004. ACM Press, pp.10–16.
- [15] Stojmenović I, Lin X. Power-aware localized routing in wireless networks. *IEEE Trans. Parall. Distrib. Syst.*, October 2001, 12(10): 1–12.
- [16] Yu Y, Govindan R, Estrin D. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. *Technical Report UCLA/CSD-TR-01-0023*, Computer Science Department, UCLA, May 2001.
- [17] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks. In *Proc. the 33rd Hawaii Int. Conf. System Sciences*, Maui, Hawaii, USA, January 2000, pp.3005–3014.
- [18] Lindsey S, Raghavendra C S. PEGASIS: Power-efficient gathering in sensor information systems. In *Proc. IEEE Aerospace Conf.*, Big Sky, Montana, USA, March 2002, pp.1125–1130.
- [19] Schurgers C, Tsiatsis V, Ganeriwal S *et al.* Topology management for sensor networks: Exploiting latency and density. In *Proc. the 3rd ACM Int. Symp. Mobile Ad Hoc Networking and Computing*, Lausanne, Switzerland, 2002, pp.135–145.
- [20] Ye W, Heidemann J, Estrin D. An energy-efficient MAC protocol for wireless sensor networks. In *Proc. IEEE INFOCOM 2002*, New York, USA, June 2002, pp.1567–1576.
- [21] Awerbuch B, Holmer D, Rubens H. The pulse protocol: Energy efficient infrastructure access. In *Proc. IEEE INFOCOM 2004*, Hong Kong, China, March 2004, pp.1467–1478.



**Jing Zhou** received her B.Eng. and M. Eng. degrees in computer science from Wuhan Technical University of Surveying and Mapping, P.R. China, in 1997 and 2000, and obtained the Ph.D. degree in computer science from the University of Southampton, UK, in 2004. From 2004 to 2006, she had been a researcher in the School of Electronics

and Computer Science at the University of Southampton. She is currently with the School of Computer and Software at the Communication University of China. Her main research interests are developing decentralized algorithms to control large scale peer-to-peer systems.



**David De Roure** graduated in mathematics with physics from the University of Southampton, UK, in 1984 and obtained the Ph.D. degree in distributed systems in 1990. He is a full professor in the School of Electronics and Computer Science at the University of Southampton, where he was a founding member of the Intelligence, Agents, Multimedia Group

and is currently head of the Grid and Pervasive Computing Group. He is Co-Director of the Open Middleware Infrastructure Institute and the Southampton Regional e-Science Centre, and Director of the Centre for Pervasive Computing in the Environment funded by the UK Department of Trade and Industry. His research interest is in the application of knowledge technologies to Grid and pervasive computing. Professor De Roure chairs the Semantic Grid Research Group in the Global Grid Forum and sits on the Grid Forum Steering Group as an Area Director. He is also a member of the W3C Advisory Committee and national committees including the e-Science Architecture Task Force and the JISC Committee for Support of Research. He is a Fellow of the British Computer Society.