Water Flow Driven Sensor Networks for Leakage and Contamination Monitoring

Amitangshu Pal and Krishna Kant
Computer and Information Sciences, Temple University, Philadelphia, PA 19122
E-mail:{amitangshu.pal,kkant}@temple.edu

Abstract- In this paper, we introduce the concept of Water flow Driven Sensor Networks for leakage and contamination monitoring in urban water distribution systems. The unique aspect of our work is that the sensor network can be deployed in the underground water network with only access to connection points (through manholes) and driven only by water harvested energy so as to avoid access to AC power or need for frequent battery changes. The main problems addressed are (a) adaptation of the network to the available energy in order to maximize leak/contamination detection, and (b) minimal artificial water circulation or leakage to improve detectability during periods of almost zero natural water flow. The paper shows, through extensive simulations, that the proposed approach can drastically reduce the leakage/contamination reporting time (more than 3 hours to ∼30 minutes), and the adaptation can reduce this circulation by ∼33% and yet enhance the collected/transmitted data by 30%.

I. Introduction

Water Distribution Systems (WDS) carry fresh water from supply sources and storage reservoirs/tanks to industrial, commercial and residential areas through a complex web of pipeline systems. However, fresh water supplies continue to dwindle and by 2025, 2/3rd of the world will experience water stress and about 25% will experience abject water scarcity [1]. While the stress on urban water systems continues to increase due to movement of population to urban areas, most of these systems are in poor shape and subject to significant amount of water leaks, seepage, and contamination [2]. In the US, most water systems are 100+ years old, particularly in large cities on the east coast. For example, a 2010 audit in Philadelphia revealed 26% water loss due to leakage and another 8% due to metering inaccuracies, water theft, and data handling and management issues [3]. A comprehensive survey in [4] shows loss percentages ranging from 15% to 35% over 36 cities in US. Europe loses more than 25% of its water to leaks, with some countries reaching 50% mark [5]. Contamination goes hand in hand with leakage due to seepage through leaks, rusted pipes, internal build ups, operational mistakes, etc. Given this state of affairs, there is a great urgency to develop ICT based solutions that can detect and localize leaks and contamination much more cheaply than mostly manual procedures followed today. A quick detection also helps to increase the working lifetime of these systems and is immensely valuable to cash-strapped water distribution utilities.

The main objective of this paper is to develop a sensor network that continuously monitors water leaks and contamination in water pipes and reports relevant data to a control station that can do the necessary analytics for detection and localization. Though conceptually straightforward, effective solutions to this problem are extremely challenging, and most proposed solutions in the literature are not very practical. The first major difficulty is that water pipes are mostly buried underground, and only accessible at connection points through manholes. Thus deploying sensors on pipes at arbitrary points, as discussed in many proposals such as PipeNet [6], MISE-PIPE [7] is infeasible or too expensive for cash strapped utilities. A related difficulty is that manholes rarely have access to AC power, and changing batteries regularly in manholes can be quite expensive. Harvesting solar energy is also impractical since the panels need to stick out of the manhole.

978-1-4799-8461-9/15/$31.00

To address these issues, in this paper we consider *water flow driven sensor networks (WDSN)* that are entirely powered by water flow via a small hydro fan unit. We use a small super-capacitor for storing the harvested energy, primarily because of the long cycle life and high charge-discharge efficiency of current super-capacitors [8]. The sensor node is assumed to be at pipe connection or valve points only, installed through the manholes. Its lower part dips into the water for energy harvesting and measurement of contamination, velocity, etc., and the upper part sports the energy storage, voltage booster, regulator and computing/communications unit. We assume that the upper part has a suitable wireless radio (e.g., WiFi) with antennas embedded on the exposed side of manhole covers (and perhaps sticking out if practical).

Due to the varying flow rate in the pipes (driven by water consumption), the availability of harvested energy varies both in spatial and temporal domains. In branch pipes, the flow rate may vary significantly and drop to near zero late at night. Thus it is essential to adapt the data collection and transmission to the available energy profile. By exploiting the highly correlated detection ability of the individual sensor nodes, we develop a dynamic sampling and transmission rate adaptation scheme based on individual node's energy budget. Furthermore, we also study the role of artificial water circulation and/or leakage mechanisms in keeping the network alive during low natural flow rates and thereby improve the contamination/leakage detection capabilities. Thus, the twin objectives of the paper are optimal rate adaptation coupled with an optimal mechanism for artificial water circulation when needed. We assume that all of the analytics are done at the central control node that has adequate computing and electric power.

The paper quantifies the advantages of our approach via extensive simulation studies using available measurement data and shows that the proposed mechanism can reduce reporting time from more than 3 hours to ∼30 minutes during late nights with only very small artificial circulation when needed. To the best of our knowledge, this is the first work on water distribution systems of its kind – one that tries to achieve optimal monitoring under the real-world practical constraints of the water networks.

The outline of the paper is as follows. Section II describes the WDSN and the artificial water circulation mechanism. Section III then develops a WDSN charging model to provide adequate energy to the sensor network during low water flow periods. Section IV-V address the problem of sensing/transmission rate adaptation based on the node's energy budget. Section VI then presents the evaluation of the scheme.

II. Water Flow Driven Sensor Network (WDSN)

We consider a large water distribution network such as the one in Fig 1 (taken from Philadelphia's water network and showing only a tiny part of the entire network). Typical water distribution system consists of *main* lines running from the reservoirs, and further divided into *sub-mains* and *branch* lines from where service connections are given to the customers. Water distribution networks are normally divided up into *District Metering Areas* or DMAs, with ability to not only measure relevant parameters such as inflows, outflows, flow head (pressure), etc., but also to control them. In dense urban areas, a DMA may consist of couple of city blocks, and we consider that as the object of study here.

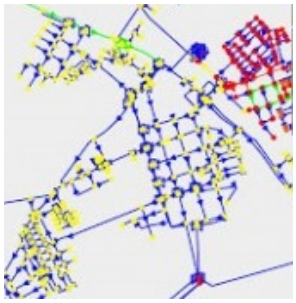


Fig. 1. A sample water distribution system.

We assume a software based leak/contamination detection scheme, where the whole DMA is modeled in a simulator [9]. The sensor nodes placed at different sections of the DMA record and report different contamination monitoring parameters along with pressure, temperature, velocity etc. These sensor readings are then compared against the simulated values at those points. If the sensor data from a node shows a wide variation from the simulator's output, a leak/contamination is suspected in the nearby regions of the sensor node.

A. Network Operation

As stated earlier, we assume that the sensing and communication nodes are deployed only at connection points and harvest energy from flowing water. We assume that the nodes are not time synchronized and use the basic *Low Power Listening (LPL)* [10] principle to conserve energy. In LPL, idle receivers run on a suitable sleep/awake duty cycle, and the senders always prepend their message with a sufficiently long preamble to ensure communication with a receiver caught sleeping. In addition, we also assume a set of strategically deployed sink nodes for data collection. These sink nodes are assumed to have a steady source of power (e.g., AC or long lasting batteries) and have a second communication interface (likely wired) to the central control node for the DMA. We assume that the sink nodes are deployed separate from the limitation of manhole locations – based on accessibility, power availability, and security considerations. All non-sink nodes collect, store and forward their sensing data and remaining energy to their nearest sink node using *single-hop, direct WiFi/long-range Zigbee communication*. While, in general, multi-hop communication may be required to cover the entire DMA, and can be enabled via LPL mechanism, we limit ourselves to single hop communication in this paper. The single hop limitation may require deploying multiple sink nodes in a DMA, but it suffices to pretend that there is a single virtual sink node for the entire DMA. The reason is that we are assuming sink nodes to be energy sufficient and always active, and thus the additional details of communication with the control center are not interesting for this paper.

The energy harvested by the normal sensor nodes in a WDSN depends on the water flow rate. We assume that each node is equipped with a suitable fan based harvester, where kinetic energy of the streaming fluid rotates the blade and generates electricity. The basic equations governing this energy conversion are well established [11]. The kinetic power (in Watts) of the moving fluid at velocity v (m/s), passing through the fan of area $\mathbb{A}$ (m^2) is given by

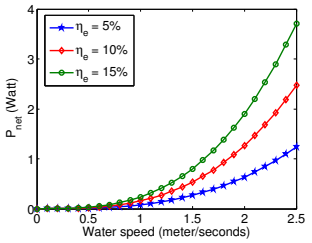


Fig. 2. P_{net} with different water velocities

$$P = \frac{\partial}{\partial t}\left(\frac{1}{2}mv^2\right) = \frac{1}{2}v^2\frac{\partial m}{\partial t} = \frac{1}{2}v^2\rho\mathbb{A}v = \frac{1}{2}\rho\mathbb{A}v^3 \qquad (1)$$

where m is the mass of the fluid and ρ is the density (1000kg/m^3 for water). Let η_{fd} denote the fluid-dynamic efficiency of the rotating body, η_{em} the electro-mechanical conversion efficiency, and η_{sc} the charging efficiency of super-capacitor. Then the net electrical power for super-capacitor charging is

$$P_{\text{net}} = \eta_{fd}\eta_{em}\eta_{sc} = \eta_e.P \qquad (2)$$

where η_e denotes the overall efficiency. Needless to say, η_e depends on a large variety of factors, and is the domain of mechanical/electrical design. Here we only assume a plausible range for η_e – usually not much better than $\tilde{1}0\%$. Fig. 2 shows the final harvested power P_{net} as a function of water flow rate with 2.5" diameter fan and $\eta_e = 5-15\%$. The most interesting aspect of this graph is that at very low water velocities, the harvested energy is effectively zero, thus requiring effective management to ensure operation during low flow periods, particularly late at night when the flow rate may stay low for hours.

B. Keeping Network Alive

The simplest approach to keep the WDSN alive is to simply choose adequate capacity super-capacitors to get through long lull periods. However, this approach not only makes the solution very expensive but also ignores an important aspect of water distribution networks: if the flow rate is very low, the leakage rate and contamination spread rates will also be very low. Thus a better idea is to *adapt* the sampling and transmission rates to the charging rate and thereby provide effective coverage without needing large energy storage. In fact, there is a sort of inherent compensation mechanism here: if a large leak develops during the lull period, the relevant sensors will automatically get charged up and become operational. Similarly, a very leaky system may always provide adequate harvestable energy, and large capacitors are wasteful. Nevertheless, it may undesirable to let the contamination monitoring frequency go down drastically during long lull periods. For this, we propose an artificial water circulation mechanism within the DMA to replenish super-capacitors.

Real water distribution systems have pumps attached at the reservoirs or tanks, and increasingly it is possible to operate them remotely from the control room [12]. The pumps can inject or extract water from the system at a certain regulated rates, which we use to generate an *occasional* artificial water circulation. Such water circulation does not entail any water loss – it is simply a circular movement among reservoirs as shown in Fig 3. Here the water injected or taken out of the system from nodes 1 and 2, and would also result in additional water flows in other nearby loops as shown. Obviously, the impact of an isolate circulation will go down rapidly as we move away from circulation area. In other words, if we want a significant artificial water flow in segments that are multiple hops away from the reservoirs and pumps (e.g., segment 5-6 in Fig 3), we would need substantial artificial flow rate which may not be possible or desirable. In those cases, we can deploy another trick – an artificial drainage of water at certain points (e.g., at node 5 or 6). This capability – operable from the control room – is also becoming increasingly available in water systems, mostly for the purposes of flushing the pipes. For obvious reasons, we want to minimize such artificial *leakage*. Artificial circulation and leakage mechanisms may also be useful without automated control, if manual action is very infrequent.

Obviously, there is a tradeoff between the capacity (and hence cost/size) of super-capacitors and the frequency and magnitude of the artificial flows created. Large super-capacitors can be charged by 1-2 significant flows during the night, but smaller ones will require many smaller flows. It is possible to define an optimization problem that determines super-capacitor sizing based on all these factors, but we do not delve into that issue for lack of space. Instead, we discuss a model for calculating the required flow rate at different connection points to provide sufficient harvesting energy for all the sensor nodes. This aspect is naturally coupled with the basic rate adaptation mechanism, which is required to minimize need for circulation.

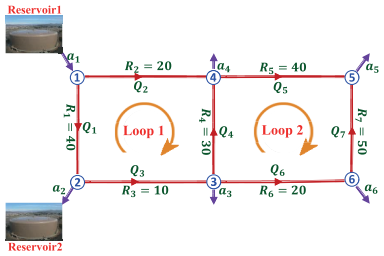


Fig. 3. Graphical representation of a simple water distribution network.

III. WDSN Artificial Water Circulation Model

We assume that all sensor nodes in the DMA are charged for a short *charging time* τ, whenever the voltage of certain number of sensor nodes drops below a threshold V_{thresh}. Let V_{ini} and V_{target} denote the initial and required target voltages, and $\mathscr{C}$ the capacitance of the super-capacitor. Then the energy stored in the super-capacitor by charging is given by:

$$P_{\text{net}}.\tau \geq \frac{1}{2}\mathscr{C}\left(V_{\text{target}}^2 - V_{\text{ini}}^2\right) \tag{3}$$

where P_{net} is the charging power in eqn (2). It follows that the required water velocity is given by:

$$v \geq \left[\mathscr{C}\left(V_{\text{target}}^2 - V_{\text{ini}}^2\right)/\tau.\rho.\mathbb{A}.\eta_e\right]^{\frac{1}{3}} = \mathbb{V}\ (\text{assume}) \tag{4}$$

Thus after charging, all nodes will have voltages of at least V_{target}. (Note that the nodes that are almost fully charged already may not take much additional charge.) The artificial water flow for charging can be generated only at *pumping points* $\mathbb{PP}$, which are the pumps associated with reservoirs/tanks. From $\mathbb{PP}$s, water can be pumped in or taken out of the system at certain regulated rates. In reality there is a maximum limit of water flow-rate that the pumping points can generate or the pipes can tolerate. We now describe three optimization problems that differ based on their design objectives.

MIN_DIFF: As the pumping points are limited, both in number as well as their pumping rates, the minimum velocity requirements of all the sensor nodes (eqn (4)) may not be met. Thus *the objective of MIN_DIFF is to minimize the sum of the differences between the required velocity and the achieved velocity at all the pipe-sections where the sensor nodes are placed*. Let a_i, Q_j, R_j denote the nodal flows at node i, pipe discharges, and pipe resistances of pipe j respectively. Let v_i and A_i denote the water velocity and area of the i-th pipe respectively, and $\mathbb{V}_j$ the minimum required velocity of the sensor node at pipe j. Let $\mathscr{V}_j^{\max}$ denote the maximum water velocity supported at pipe segment j, and $a_i^{\max}$ the maximum flow supported at pumping point i. Let $\mathbb{C}$ and $\mathbb{L}$ denote the number of connection points and loops in the distribution system respectively, $\mathbb{P}$ the set of all pipes, and $\mathbb{S}$ the subset where sensor nodes are placed. Then:

$$\begin{aligned}
\textbf{Minimize}\quad & \textstyle\sum_{j\in\mathbb{S}} \max\left(0, \mathbb{V}_j - v_j\right) \\
\text{subject to}\quad & v_j = \frac{|Q_j|}{A_j} \leq \mathscr{V}_j^{\max}, \quad \forall j \in \mathbb{P} \\
& \pm a_i \pm \sum_{\substack{\text{pipe } j \\ \text{connected to } i}} Q_j = 0, \quad \forall i = 1, 2, ..., \mathbb{C} \\
& \pm \sum_{\substack{\text{pipe } j \\ \in L_l}} R_j Q_j^2 = 0, \quad \forall L_l = L_1, L_2, ..., L_{\mathbb{L}} \\
& |a_i| \leq a_i^{\max} \quad \forall i \in \mathbb{PP}
\end{aligned} \tag{5}$$

We have used the Darcy-Weisbach formula in eqn (5) to calculate the frictional head loss. Among the non-pumping points, if there exists some background flow at some connection point i, a_i is assigned to that background flow. Otherwise a_i is assumed to be zero, at the non-pumping points. These background flows can be estimated from the water flow measurements reported by the sensor nodes.

In eqn (5) the first set of constraints ensures that the water velocity through a pipe segment j (which is equal to its volumetric flow rate/discharge divided by the cross-sectional area of the pipe segment) is less than the maximum water velocity $\mathscr{V}_j^{\max}$ that the pipe segment can support. The second set of constraints are the *node-flow continuity relationships* that ensure that the sum of the inflows and outflows at all connection points are zero. The third set of constraints are the *loop-head loss relationships* that state that the sum of head losses in pipes forming a loop is zero. $\pm$ sign is used in these two constraints to take into account the direction of the water flow assumed. The fourth constraint states that the water-flow rate at all the pumping point i is less than some maximum threshold $a_i^{\max}$. Note that when the objective value of **MIN_DIFF** is zero, the pumping points can satisfy the velocity requirements of all the sensor nodes.

MIN_PUMPING: If the velocity requirements of all sensor nodes can be satisfied, we would like to pump water in such way that the amount of water pumped is minimized. Note

that MIN_PUMPING is solved only if the result of MIN_DIFF gives zero objective value. Given the notations and explanation of MIN_DIFF, the following formulation should be clear and is not explained further:

$$\textbf{Minimize} \quad \textstyle\sum_{i\in\mathbb{PP}} |a_i|$$

$$\begin{aligned}\text{subject to} \quad & v_j = \frac{|Q_j|}{A_j} \leq \mathscr{V}_j^{\max}, \quad \forall j \in \mathbb{P} \\ & v_k \geq \mathbb{V}_k, \quad \forall k \in \mathbb{S} \\ & \pm a_i \pm \sum_{\substack{\text{pipe } j \\ \text{connected to } i}} Q_j = 0, \quad \forall i = 1, 2, ..., \mathbb{C} \\ & \pm \sum_{\substack{\text{pipe } j \\ \in L_l}} R_j Q_j^2 = 0, \quad \forall L_l = L_1, L_2, ..., L_{\mathbb{L}} \\ & |a_i| \leq a_i^{\max} \quad \forall i \in \mathbb{PP} \end{aligned} \tag{6}$$

MIN_DISCHARGE: This problem is a variant of MIN_PUMPING problem where we assume the existence of a set of *discharge* (or deliberate leakage) points $\mathbb{LP}$ as well, from where water can be deliberately leaked at different regulated rates. As the number of pumping points is limited, the idea is to discharge/leak some amount of water (only for the charging time) to generate certain water-flow at few pipe-sections that keeps the sensor nodes running. Since the discharge wastes water, we want to minimize it. This can be modeled as follows:

$$\textbf{Minimize} \quad \textstyle\sum_{i\in\mathbb{LP}} |a_i|$$

$$\begin{aligned}\text{subject to} \quad & v_j = \frac{|Q_j|}{A_j} \leq \mathscr{V}_j^{\max}, \quad \forall j \in \mathbb{P} \\ & v_k \geq \mathbb{V}_k, \quad \forall k \in \mathbb{S} \\ & \pm a_i \pm \sum_{\substack{\text{pipe } j \\ \text{connected to } i}} Q_j = 0, \quad \forall i = 1, 2, ..., \mathbb{C} \\ & \pm \sum_{\substack{\text{pipe } j \\ \in L_l}} R_j Q_j^2 = 0, \quad \forall L_l = L_1, L_2, ..., L_{\mathbb{L}} \\ & |a_i| \leq a_i^{\max} \quad \forall i \in \mathbb{PP} \cup \mathbb{LP} \end{aligned} \tag{7}$$

In the above formulation, the zero objective value means that without any water discharge (waste), the pumping points can satisfy the node's demands. In that case the secondary objective of maintaining a minimum flow-rate at the pumping points, can be achieved by again solving the **MIN_PUMPING** problem as before.

We illustrate these problems using the simple example in Fig. 3. Let us assume that the water is pumped in from reservoir 1 and pumped out from reservoir 2. Let nodes 3-6 be discharge points, with discharge rates of a_3-a_6. Thus the *node-flow continuity relationships* and the *loop-head loss relationships* are as follows:

$$\begin{aligned} & a_1 - Q_1 - Q_2 = 0 && Q_1 - a_2 - Q_3 = 0 \\ & Q_3 - a_3 - Q_4 - Q_6 = 0 && Q_4 + Q_2 - a_4 - Q_5 = 0 \\ & Q_5 + Q_7 - a_5 = 0 && Q_6 - a_6 - Q_7 = 0 \\ & R_2Q_2^2 - R_4Q_4^2 - R_3Q_3^2 - R_1Q_1^2 = 0 \\ & R_5Q_5^2 - R_7Q_7^2 - R_6Q_6^2 + R_4Q_4^2 = 0 \end{aligned} \tag{8}$$

We solve the above optimization problems for this case using **AMPL**, which is a modeling language for solving large-scale optimization problems [13]. The necessary parameters are listed in Table I. The maximum pumping capacity $a^{\max}$ (assumed to be 0.069 m^3/s in Table I) corresponds to a flow rate of 7 ft/sec in a pipe with 8 inches diameter. Fig. 4 shows the variation of $\sum_{j\in\mathbb{P}} \max(0, \mathbb{V}_j - v_j)$ with different charging

TABLE I. PARAMETERS USED

Var	Values	Var	Values	Var	Values
ρ	1000 kg/m^3	Fan diameter	2.5 inch	Pipe diameter	8 inch
$a^{\max}$	0.069 m^3/s	$\mathscr{V}_j^{\max}$	7 ft/sec	$\mathscr{C}$	25 F
V_{ini}	0.9 V	V_{target}	1.5 - 2 V	η_e	10%

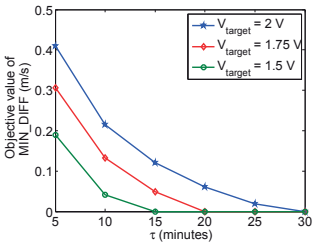


Fig. 4. Objective value of MIN_DIFF vs charging time τ.

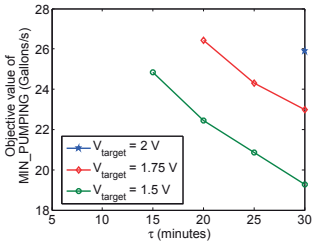


Fig. 5. Objective value of MIN_PUMPING vs charging time τ.

times. As expected, the objective value decreases with the increase in charging time and with reduced target voltage V_{target}. Fig. 5 shows the total amount of water pumped in and out through the $\mathbb{PP}$s, with the variation of charging time. In Fig. 5, initially the optimization problem MIN_PUMPING is infeasible, so there are no points in the graph. When the objective value of MIN_DIFF is zero, MIN_PUMPING starts giving feasible solutions, which is also a decreasing function of charging time.

Fig. 6 shows the total amount of discharge with different charging times, where the maximum discharge rate is assumed to be same as the maximum pumping rates. Note that Fig. 4 and Fig. 6 show a strong similarity, this is because whenever there is a non-zero difference between the required and achieved velocity in MIN_DIFF, there needs to be some non-zero discharges in MIN_DISCHARGE. Also the total discharges decrease with increasing charging time, as the sensor nodes get more time to replenish their super-capacitors.

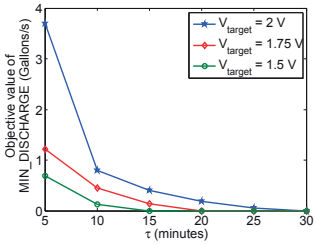


Fig. 6. Objective value of MIN_DISCHARGE vs charging time τ.

IV. SAMPLING AND TRANSMISSION RATE ADAPTATION

As discussed earlier, adaptation of the measurement/transmission activity to current state of the charge in super-capacitors is crucial for maintaining maximal coverage of the leakage/contamination detection activity. This is true even with artificial water circulation, since sensor nodes in certain segments of the water network may be difficult to charge effectively.

In general, a set of sensor nodes in a vicinity may have significant dependency with respect to water flows and hence their chargeability and leakage/contamination detection performance. We can consider these nodes as forming a *coalition* in the game theoretic sense which can be exploited for improved performance. Coalition can be formed by simulating leaks/contaminations at different pipe sections, using any commercial simulator such as Water-GEMS [9] and by looking at the inter-dependencies among the detection abilities of the individual nodes, i.e. if there is a leak/contamination at any

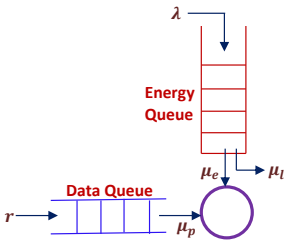


Fig. 7. Block diagram of individual node's data queue and energy queue.

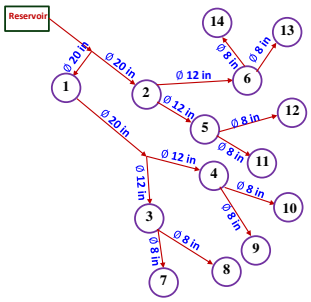


Fig. 8. Simulation topology, red arrows show the direction of water-flow.

pipe section in a coalition, at least few sensor nodes are able to detect it. The coalition members can collaboratively adapt their sampling rates based on the individual node's energy availability, i.e. the low sampling rate of the sensor nodes with low harvested energy is compensated by the higher sampling rate of the nodes with higher energy. Such a mechanism is quite different from the individual node-based rate/energy allocation schemes discussed in the literature [14], [15].

We assume that time is divided into *intervals* of T time units. The sampling rates are updated periodically in every interval as follows. All nodes keep track of their average harvested energy in each interval. Based on their historical energy profiles, they predict their expected harvested energy for the next interval, which is then used by the sensor nodes to calculate their maximum possible sampling rate, as described in section IV-A. The maximum possible sampling rate is broadcast using *beacon* messages. The sink uses these rates to compute the optimal sampling rates of individual nodes and broadcasts by sending beacons, as described in section IV-B.

A. Predicting Energy Harvesting

Fig. 7 shows the conceptual model for energy harvesting. Available energy is stored in the *energy queue (EQ)* which is a super-capacitor in our case. The sampled values are stored in a volatile RAM, which we call *data queue (DQ)*. While transmitting packets, a sensor node takes ℓ items from the RAM with $\ell_{\min} \leq \ell \leq \ell_{\max}$. Here $\ell_{\min}$ is the minimum number of samples that a node will accumulate before transmitting if it has enough energy to do immediate transmission. On the other hand, if the node is low on energy, it will continue sampling and storing samples in the RAM (if possible). When the next energy burst arrives, it will transmit all accumulated samples up to the limit of $\ell_{\max}$.

Each sensor node estimates the energy arrival in its super-capacitor in periodic intervals of T using a normalized least mean square (NLMS) adaptive filter. In NLMS filter, historical harvested energy profile is stored in the vector E_{t-1}. Based on this profile, predicted harvested energy for the next interval λ_t is calculated by a dot product between E_{t-1} and the coefficients of the adaptive filter $\mathbb{W}_{t-1}$ using $\lambda_t = E_{t-1}\mathbb{W}_{t-1}$ and the error e_t is recorded. The filter coefficient is then modified as:

$$\mathbb{W}_t = \mathbb{W}_{t-1} + \frac{s.e_t.E_{t-1}}{1+|E_{t-1}|^2} \tag{9}$$

where s is the step size of the filter. The super-capacitor leakage power and average power consumption due to different operations (sensing/transmission/reception etc) are assumed to be μ_l and μ_e respectively. The average power consumption μ_e needs to be adapted based on the energy availability to maintain the energy conservation, i.e.

$$A_e + \lambda_t - (\mu_e + \mu_l)T \geq 0 \quad \therefore \mu_e \leq \frac{A_e + \lambda_t}{T} - \mu_l \tag{10}$$

where A_e is the amount of available energy at the beginning of that interval t.

The sampled values are stored in the DQ with an arrival rate of r, while the packet transmission rate is μ_p. Note that λ_t and μ_e are expressed in units of energy, whereas r is expressed in number of samples. We calculate the maximum sampling rate that the sensor node can support in the next time interval, without DQ buffer overflow. Assume that at the time of computing the maximum sampling rate, the number of packets waiting in the DQ is $\mathcal{N}$. The DQ capacity is assumed to be C. To maintain the energy budget, $\mu_e = \mathcal{A}.r + \mathcal{B}.\mu_p + \mathcal{C}$, where $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are constants that capture the power consumption due to sensing, transmission and other operations (beacon transmission/reception, processing etc) respectively. To avoid DQ buffer overflow

$$\begin{aligned} &\mathcal{N} + (r - \ell_m.\mu_p).T \leq C \\ &\therefore r \leq \frac{C-\mathcal{N}}{T} + \ell_m.\mu_p = \frac{C-\mathcal{N}}{T} + \ell_m \frac{\mu_e - \mathcal{A}.r - \mathcal{C}}{\mathcal{B}} \\ &\therefore r \leq \frac{\frac{C-\mathcal{N}}{T} + \ell_m \frac{\mu_e - \mathcal{C}}{\mathcal{B}}}{1 + \ell_m.\frac{\mathcal{A}}{\mathcal{B}}} = \mathbb{R} \quad \text{(assume)} \end{aligned} \tag{11}$$

which gives the upper limit on the sampling rate. All sensor nodes periodically calculate their maximum sampling rate $\mathbb{R}$ and broadcast them in their beacon messages, which is used by the sink to determine the sampling rates of all the individual sensor nodes.

B. Computing Optimal Sampling Rate

Upon receiving the maximum sampling rate $\mathbb{R}$ from all the sensor nodes, the sink formulates the sampling rate adaptation problem to maximize a certain utility function, under the required energy constraints. Suppose that there are N nodes in a coalition. As the detection abilities of the sensor nodes in a coalition are highly correlated, the sensor nodes in a coalition can share the data sampling task among themselves for reduced energy consumption, based on their available harvested energy. We define the utility of a node i by considering two factors

- The sensing rate r_i. As r_i increases the number of sampled points increases and so does the utility.

- In a WDSN, main lines are generally more important than branch lines, as water from the main lines are distributed to different sub-mains and branches. Thus a sensor node placed in a main line is considered to be more important than sensors placed in branches. Thus we assign a relative weight α_i to the sampled data of node i, based on its position in the WDSN.

Beyond the distributional hierarchy, there may be other considerations in assigning the weights α_i, as determined by water system personnel. For example, the water pressure often varies significantly within a DMA, and nodes in higher pressure area can be given higher weights because of greater water loss and more potential damage due to leaks there. For contamination monitoring, one can assign weights to the nodes that are close to the reservoirs, because any contamination

close to the reservoir needs to be detected sooner, to avoid its spread. Higher weights can also be assigned to older and more damage prone pipes. Also in a coalition, the detection abilities of certain sensor nodes may be higher compared to others, thus those nodes can be assigned higher weights.

Considering these factors, the **weighted proportional fairness** within a coalition can be achieved by modeling the utility function of node i as $U_i(r_i) = \alpha_i.\log(r_i)$, where α_i is the normalized weight. Our objective is to maximize the overall utility of the coalition, i.e. $\sum_{i=1}^{N} U_i(r_i)$, after satisfying the energy budget of individual nodes. We also assume that the sink places an upper limit of M samples/interval from a coalition, to avoid redundant sampling, i.e. $\sum_{i=1}^{N} r_i.T \leq M$, or $\sum_{i=1}^{N} r_i \leq \frac{M}{T} = \mathbb{M}$. Intuitively we can think that the sensor nodes in a coalition work as a *single virtual sensor node* that senses and reports at a maximum rate of M samples/interval. M is basically a controlling parameter that controls the overall sampling rate of the coalition, i.e. if the sink wants to receive the samples more frequently, it increases M and vice versa. Thus the optimization problem can be written as

$$\begin{aligned} \text{Maximize} \quad & \sum_{i=1}^{N} U_i(r_i) \\ \text{subject to} \quad & \sum_{i=1}^{N} r_i \leq \mathbb{M}, \quad r_i \leq \mathbb{R}_i, \forall i, \quad r_i \geq 0, \forall i \end{aligned} \tag{12}$$

where $r_i \leq \mathbb{R}_i$ is the maximum sampling rate constraint (MSRC) obtained from equation (11). As log is a concave function, this problem is a convex optimization problem, that can be solved centrally by solving the corresponding Lagrangian and KKT conditions. We propose an algorithm to solve this problem, which is presented in the section V.

Algorithm 1 Collaborative Adaptive Rate Allocation scheme (CARA)

1: INPUT : Maximum sampling rate $\mathbb{R}_i$, utility weights α_i and $\mathbb{M}$.
2: OUTPUT : Sampling rates r_i $\forall i \in \{1, 2, ..., N\}$.
3: A=$\{\phi\}$; U=$\{1, ..., N\}$;
4: **for** each node $i = \{1, 2, ..., N\}$ **do**
5: $r_i = \frac{\alpha_i}{\sum_{i \in U} \alpha_i} \mathbb{M}$;
6: diff[i] = $\mathbb{R}_i - r_i$;
7: **end for**
8: **for** each node $k = \{1, 2, ..., N\}$ **do**
9: Sort node $\in U$ in increasing order of diff[k];
10: Put them in order in list L;
11: j = L[0];
12: **if** diff[j] < 0 **then**
13: $r_j = \mathbb{R}_j$; $A = A \cup j$; $U = U \backslash j$;
14: **for** each node $i = \{1, 2, ..., N\}$ AND $i \in U$ **do**
15: $r_i = r_i + \frac{\alpha_i}{\sum_{i \in U} \alpha_i}$.abs(diff[$j$]);
16: diff[i] = $\mathbb{R}_i - r_i$;
17: **end for**
18: diff[j] = 0;
19: **else**
20: EXIT
21: **end if**
22: **end for**
23: return r_i $\forall i$

V. Proposed Rate Adaptation Scheme CARA

Based on the steps described above, we now describe our proposed *Collaborative and Adaptive Rate Allocation (CARA)* scheme, as shown in Algorithm 1. In this algorithm, the sink maintains two sets of nodes: unassigned U and assigned A. Initially, all nodes belong to set U, but are transferred to set A as rates are assigned for them. The sink first assigns the sampling rates to each sensor i as $r_i = \frac{\alpha_i}{\sum_{i \in U} \alpha_i} \mathbb{M}$ (line 4-7). The difference between assigned sampling rate r_i and maximum sampling rate $\mathbb{R}_i$ is stored in diff[i]. After this sampling rate assignment, if the MSRC (obtained from equation (11)) is violated for any node j, then diff[j] < 0. For those nodes, the sink assigns their rates as their maximum rate $\mathbb{R}_j$ (line 11) and divides the diff[j] fairly among other nodes (line 14-17). This process is continued until the MSRC is satisfied for all the nodes. The calculated sampling rates are broadcasted to all the sensor nodes.

Optimality of the proposed scheme: The proposed scheme assigns the sampling rate fairly to all the nodes based on their weighted utilities considering the energy constraints. We prove that the proposed scheme is optimal under the given assumptions. We first propose and prove Lemma 1 and Lemma 2 as follows.

Lemma 1. *The solution of the optimization problem:*

$$\textbf{Max} \sum_{i=1}^{N} U_i(r_i) = \sum_{i=1}^{N} \alpha_i.\log(r_i) \quad \text{s.t.} \sum_{i=1}^{N} r_i \leq \mathbb{M}, \; r_i \geq 0, \forall i \tag{13}$$

is $r_i = \frac{\alpha_i}{\sum_i \alpha_i} \mathbb{M}$, $\forall \; i = \{1, 2, ..., N\}$.

Proof: Clearly r_i cannot be zero for any i. This is because making r_i = 0 makes the objective value $-\infty$. Thus the last constraint is inactive. Then by solving the KKT conditions of problem (13), we obtain $r_i = \frac{\alpha_i}{\sum_i \alpha_i} \mathbb{M}$ (details are skipped due to space limitations). ■

Lemma 2. *If the sampling rate of a node is reduced by an amount δ from the optimal rate in problem (13), and divided among others proportionately, then the overall objective function is a decreasing function of δ.*

Proof: Let us assume that for any node j, we assign an amount $\frac{\alpha_j}{\sum_i \alpha_i} \mathbb{M} - \delta$ and divide δ among all others proportionately, so that all nodes $i \neq j$ are assigned a rate of $\frac{\alpha_i}{\sum_i \alpha_i} \mathbb{M} + \frac{\alpha_i}{\sum_{i \neq j} \alpha_i} \delta$. Assume $\Gamma = \sum_{i=1}^{N} \alpha_i$ and $\Delta = \sum_{i \neq j} \alpha_i$. If the new objective function if $F(.)$ then we can show that

$$\begin{aligned} F &= \alpha_j \log\left(\frac{\alpha_j}{\Gamma}\mathbb{M} - \delta\right) + \sum_{i \neq j} \alpha_i \log\left(\frac{\alpha_i}{\Gamma}\mathbb{M} + \frac{\alpha_i}{\Delta}\delta\right) \\ \frac{\partial F}{\partial \delta} &= -\frac{\alpha_j}{\frac{\alpha_j}{\Gamma}\mathbb{M} - \delta} + \sum_{i \neq j} \frac{\alpha_i}{\Delta\left(\frac{\mathbb{M}}{\Gamma} + \frac{\delta}{\Delta}\right)} < 0 \end{aligned} \tag{14}$$

Thus $F(.)$ is a strictly decreasing function of δ. ■

Theorem 3. *The proposed CARA algorithm gives optimal solution for problem (12).*

Proof: From Lemma 1, we get $r_i = \frac{\alpha_i}{\Gamma} \mathbb{M}$ for the optimization problem (13). Now we introduce the maximum sampling rate constraint (MSRC) $r_i \leq \mathbb{R}_i$ in problem (13). Suppose $r_j = \frac{\alpha_j}{\Gamma} \mathbb{M}$ violates the MSRC of node j, i.e. $r_j > \mathbb{R}_j$ and diff[j] = $\mathbb{R}_j - r_j = \mathbb{R}_j - \frac{\alpha_j}{\Gamma} \mathbb{M}$.

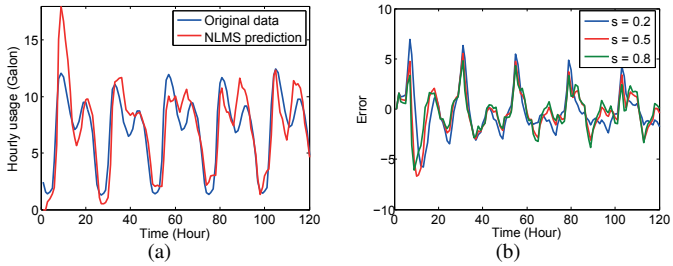
We consider this problem in two steps. In the first step, we divide a total sampling rate of $\tilde{\mathbb{M}} = \frac{\Gamma}{\alpha_j} \mathbb{R}_j$ over N nodes. Then using Lemma 1, $r_j = \frac{\alpha_j}{\Gamma} . \frac{\Gamma}{\alpha_j} \mathbb{R}_j = \mathbb{R}_j$. Thus node j's MSRC is satisfied. For any other node i, $\hat{r_i} = \frac{\alpha_i}{\Gamma} . \frac{\Gamma}{\alpha_j} . \mathbb{R}_i = \frac{\alpha_i}{\alpha_j} \mathbb{R}_j$. At this point node j's utility cannot be improved any further by changing r_j (reducing r_j results in degradation of the overall objective as shown in Lemma 2).

Now we introduce $\mathbb{M} - \tilde{\mathbb{M}}$ amount of additional sampling rates to this system. Clearly node j cannot be assigned more rates. Thus we assign $\mathbb{M} - \tilde{\mathbb{M}}$ among all $i \neq j$ fairly. Using Lemma 1, the new rates of all $i \neq j$ are

$$\begin{aligned} r_i^{\text{new}} =& \hat{r_i} + \frac{\alpha_i}{\Delta}.\left(\mathbb{M} - \tilde{\mathbb{M}}\right) = \frac{\alpha_i}{\alpha_j}\mathbb{R}_j + \frac{\alpha_i}{\Delta}.\mathbb{M} - \frac{\alpha_i}{\Delta}.\frac{\Gamma}{\alpha_j}.\mathbb{R}_j \\ =& \frac{\alpha_i}{\Gamma}.\mathbb{M} + \frac{\alpha_i}{\Delta}\left(\frac{\alpha_j}{\Gamma}.\mathbb{M} - \mathbb{R}_j\right) = r_i + \frac{\alpha_i}{\sum_{i \neq j} \alpha_i}.\text{abs (diff}[j]) \end{aligned} \quad (15)$$

which is same as the rate assigned according to Algorithm 1 (line 15). At this stage, if any other nodes violate the MSRC constraint, we do similar operation till the MSRC constraint is fulfilled for all the nodes. Thus the optimal sampling rates are obtained for all the nodes. ■

Complexity analysis: Algorithm 1 runs in $O(N^2 \log N)$ time when the number of nodes is N. We first focus on one iteration of the outer for loop of line 8-22. The sorting operation in line 9 takes $O(N \log N)$ time. The inner for loop of lines 14-17 take additional $O(N)$ time. As the outer loop runs $O(N)$ times, hence the time complexity of the whole scheme is $O(N^2 \log N + N^2) = O(N^2 \log N)$.



(a) (b)

Fig. 9. (a) Hourly water usage of a single-family home for five days (s = 0.8). (b) Convergence of the forward predictor with different s.

VI. Simulation Results

Ideally, the evaluation of the scheme should be done with a real water distribution network, however, this is simply not possible in practice. Water distribution companies are generally not even willing to share the data they already collect, much less providing access to their distribution systems. There are real challenges in putting together a realistic network in the lab (e.g., access to large volume water supply, reservoirs, energy harvesters, etc.). As a result, the evaluation in this paper is largely based on simulations that account for the water flow physics [16] and use parameters obtained from characterization of real water distribution systems.

We study the proposed rate adaptation scheme in *Castalia* [17], which is an application-level simulator for wireless sensor network based on OMNeT++. The simulated system topology along with the pipe diameters are shown in Fig. 8. Water from the reservoir comes to nodes 1 and 2 (first level nodes), distributed to nodes 3-6 (second level nodes), and then to 7-14 (third level nodes). Each node has the fan for energy harvesting, a super-capacitor, water sensors, a small computer, and WiFi radio. The cross-sectional area of the fans are chosen as $\frac{1}{16}$th of the pipe cross section, to avoid blocking the normal water flow. For simulations, τ is assumed to be 1 minute, which is much less than in section III . The difference can be attributed to the fact we now have successively smaller pipe diameters (going from levels 1 to 2 to 3) which increases water velocity and helps with respect to charging of nodes. Due to this structure, it is also reasonable to assume that all nodes fall into a single coalition.

We model the harvested energy arrival from water-flow based on the average water usage pattern, taken from [18], and shown in Fig. 9(a) for a typical single-family home over five days. The total daily usage is 169 $\pm$ 10.6 gallons. Reference [19] reports the maximum water velocity in real systems as 7.5 ft/sec. We conservatively assume that for the third level nodes have a water-velocity of 5.0 ft/sec at peak hours and compute those for other two layers using the flow continuity relationships. We also calculate water velocities and the *available energy* at other times based on the usage pattern and variation. Fig. 9(b) shows the NLMS (normalized least mean square) filter predictor of the *available energy* for five days. We use s as 0.8.

The sink node broadcasts the assigned rates every $T = 1$ hr (interval time), chosen such that the harvested energy does not change significantly within the interval time. The beacon interval of the sensor nodes is assumed to be 30 minutes. We assume that the nodes use asynchronous Low Power Listening that makes them sleep most of the time and wake-up periodically to check the channel activity. The power consumption in each node is represented as:

$$P_{\text{node}} = \frac{P_{Bt}T_{Bt}}{T_B} + \mathscr{M}.P_{Dt}T_{Dt} + \mathscr{N}.P_{Br}T_{Br} + \mathscr{S}.P_sT_s + \mathscr{P}.P_PT_P$$

where P_x and T_x represent the power consumption and the duration, respectively, of the event x; and T_B represents the beacon interval. Transmission/reception of beacons is denoted by B_t/B_r, data transmit/receive is denoted by D_t/D_r, and processing and sensing are denoted as P and S, respectively. $\mathscr{M}$, $\mathscr{N}$ and $\mathscr{S}$ are the number of data transmission, beacon reception and data sampling respectively. $\mathscr{P}$ represents the number of times that a node wakes-up per second to check if the channel is busy, and is set to 8 in our application.

We assume DQ capacity as 20 samples, $\ell_{\min}$, $\ell_{\max}$ as 3, 5 respectively, $\mathbb{M}$ as 840 samples/hr (i.e. 1 sample/min per node on average), and harvesting efficiency $\eta_e = 10\%$. In reality η_e itself is dependent on flow velocity and load, but for simplicity we keep it fixed at 10% for our simulations. The super-capacitor is assumed to be of 25Farad @2.7V with an initial voltage of 2.0V for all nodes. The super-capacitor leakage power is calculated as $P_0.\exp(a.V_c)$ [8], where V_c is the super-capacitor voltage and P_0 and a are constants obtained from best-fitting the experimentally obtained results, and are $P_0 = 2.572e^{-17}$ and $a = 11.982$ respectively. The DC-DC converter efficiency (in between the super-capacitor and the sensor node) is assumed to be 75% [8]. The sampling/transmission is stopped, whenever the capacitor voltage goes below 0.9 V, which is considered as very low voltage.

We assume α at levels 1, 2, and 3 in the ratio 4, 2, and 1 to reflect the fact that detection at higher levels of the distribution network is much more important than at lower levels. We use the following policy for water circulation: if 50% of the nodes go below the threshold voltage of V_{thresh} = 0.9 V, water is pumped in through the pipes of nodes 1-2 to boost the node voltages to V_{target} = 1.5 V (or higher) at all nodes. However a sampling frequency based pumping policy can also be adopted, i.e. whenever the overall sampling frequency goes below a limit, the pumping starts. We model both our scheme, i.e., CARA, and the simpler non-adaptive scheme called *equal rate allocation (ERA)* which assigns same sampling rate to all nodes. We use nodes 1, 3 and 7 to show the characteristics

of first, second and third level nodes respectively. We run the simulation for 24 hours. Parameters used for simulations are listed in Table II.

TABLE II. SIMULATION PARAMETERS

Var	Values	Var	Values	Var	Values	Var	Values
P_{Bt}	1000 mW	T_{Bt}	140 ms	P_{Br}	200 mW	T_{Br}	140 ms
P_{Dt}	1000 mW	T_{Dt}	140 ms	P_{Dr}	200 mW	T_{Dr}	140 ms
P_P	200 mW	T_P	3 ms	P_S	495 mW	T_S	400 ms

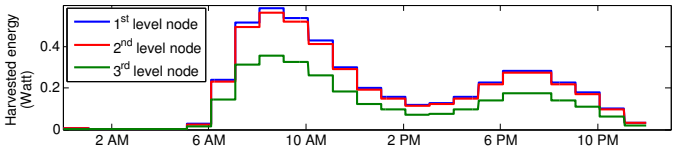


Fig. 10. Mean energy harvested over time for different nodes.

Fig. 10 shows the mean energy profile of nodes from normal water flow over 24 hrs at levels 1, 2, and 3. To model fine-grain harvesting variations due to water flow turbulence, the actual energy arrival is modeled as uniformly distributed around the mean. The energy harvested depends on the water velocity and fan diameters. In this example, the water velocity increases at lower levels but the fan diameter decreases, thereby resulting in the behavior shown.

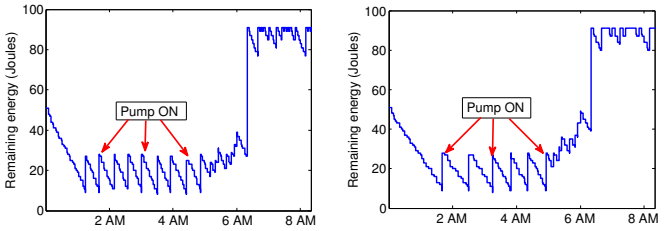


Fig. 11. Remaining energy over time for (a) ERA scheme, (b) CARA scheme.

Benefits of Energy Adaptation: We compare our rate adaptation scheme with ERA in Figs. 11(a) and (b), which show the remaining energy of node 7 during the crucial night hours. It is seen that without adaptation, the pump is on frequently because the sensor node continues to sample fast and dies more often. However, with adaptation, both the sampling rate and hence the pumping rate slow down. In particular, CARA effectively reduces the pumping frequency by 33% as shown in Fig. 12(b). Also by dynamically adapting the sampling rates to maximize the system utility, CARA achieves 35% of higher *information* measure (defined as the product of the number of samples and their relative weights) without any water circulation and 30% in presence of artificial circulation, compared to ERA as seen from Fig. 12(a).

CARA can further reduce the pumping frequency by exploiting the mutually inter-dependent detection abilities among the members of a coalition. In particular, Fig. 12(b) also shows the pumped water amount for CARA under the policy that pumping is done only when all sensor nodes within a coalition die. This policy reduces pumping frequency by another 20%.

Benefits of artificial water circulation: Fig. 13(a) and (b) show the sampling schedules of nodes 1, 3 and 7 without and with artificial water circulation. It is seen that the artificial circulation at the night drastically improves the monitoring. It is also seen that the sampling rate of the higher level nodes are higher, because of higher energy availability and higher chosen utility weight (i.e., α). This clearly shows the adaptive nature

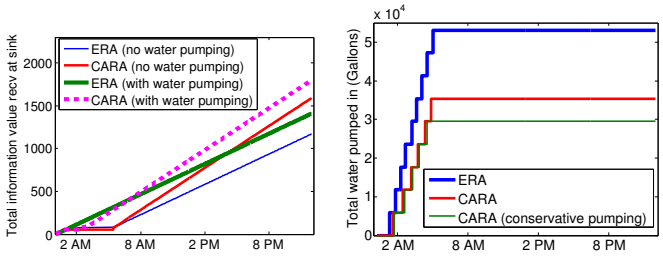


Fig. 12. (a) Comparison of total effective information received at the sink. (b) Comparison of total water pumped into the system.

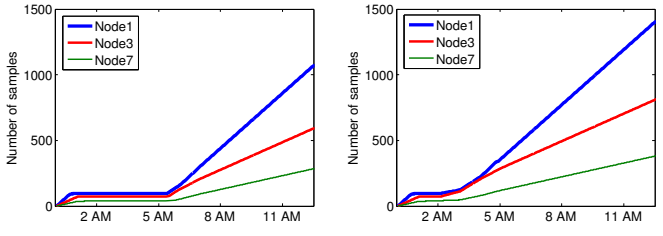


Fig. 13. Number of samples with time, (a) without water circulation (b) with water circulation.

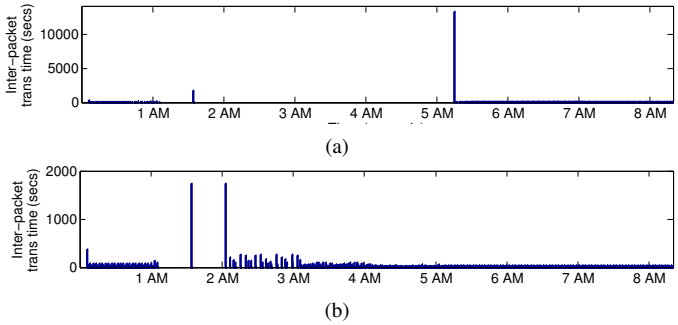


Fig. 14. Event reporting time (a) without and (b) with water circulation

of CARA based on the individual nodes relative weights and energy availability.

From Fig. 14(a) we can observe that without water circulation, it takes more than 3 hours for the system to notify the WDS administrator of small leaks/contamination spread. However, a small amount of artificial circulation reduces this to about 30 minutes as seen from Fig. 14(b). By providing the extra sampling capability at night, the water circulation procedure improves the effective *information* measure by 14% as observed from Fig. 12(a). Also in this example, the amount of extra water circulation during night hours is only a small fraction ($< 0.7\%$) of the total water-flow throughout the day as derived from Fig. 12(b).

VII. RELATED WORK

Wireless sensor networks for pipeline monitoring is well-researched, but most of them require manual access to the pipes and thus not very useful for much of practical WDS. For example, PipeNet [6], MISE-PIPE [7] involve sensor deployment along pipe length on the outside whereas SPAMMS [20] requires RFID tags painted inside of the pipe. In [21] the authors propose a self maneuvering robot going through the pipes and monitors the leaks. Authors in [22] consider flowing acoustic sensors that communicate with fixed relay nodes deployed along the pipe. Contrary to these approaches, we focus on techniques that do not require any ground digging or even following of buried pipelines from above-ground, which itself

may be impractical. Different mathematical models have been explored for leak modeling as well by considering them as demands that depend on the pressure [23], [24]. WaterGEMS is a commercial package that incorporates many such features for designing and configuring WDS [9].

Energy management in sensor network is also heavily researched. Control of sleep/wakeup cycle is a standard technique that is explored in several MAC proposals [25]. Other techniques for reducing energy consumption include data compression and source coding [26], transmit power control [27], multiple channel assignment [28] etc. While these proposals are mainly motivated towards maximizing the life-time of the sensor network, our objective in this paper is to schedule the operations according to the energy harvesting opportunities and adapt them to the energy availability that varies dynamically while maximizing the collection of most useful samples. In this regard, some relevant papers are [29] and [30], where the authors propose fair rate adaptation for interference or congestion control; however, they do not consider adaptation for meeting individual node's energy budget. Authors in [14] and [15] propose energy aware rate adaptation schemes using dual decomposition in a distributed manner, that can incur high control overhead and long running time, which make their schemes impractical especially in the context of resource constrained sensor networks. In contrast, our technique is a *collaborative* rate adaptation that exploits correlated detection of a "coalition" of sensor nodes and is computed in a *centralized* manner to avoid the overhead of distributed computations. Such a scheme can be used in other energy harvesting environments where the sensor nodes have correlated event detection or sensing capabilities.

VIII. Conclusions

In this paper, we explored water flow driven sensor network that can be practically deployed in real water distribution networks without substantial expense and continued maintenance. We showed that the scheme can significantly reduce the leak/contamination detection time during periods of low water flow, while minimizing the need for artificial water flow to keep the network alive. We also motivated the advantage of collaborative sampling within a coalition, and we plan to explore this aspect along with the optimal coalition formation in more detail in the future. Another important concern of WDSN is its security [31], [32], even if the sensor nodes are physically secured. An adversary can eavesdrop to acquire secure network information or maliciously inject tampered packets into the network to produce false leak/contamination alarms. To make the WDSN dysfunctional, an attacker can also inject jamming signals, which impedes the wireless communications and at the same time depletes the receiver's super-capacitors. Making the resource constrained WDSN resistant to these security concerns is one of our future research endeavors. We also plan to conduct more comprehensive simulations that more directly include the effect of flow turbulence and obstructions/build ups within the pipes on flows and energy harvesting. Other than simulations, we also plan to do real experiments, to the extent such experiments are feasible based on the available monitor and energy harvesting products and tools.

References

[1] D. Seckler, R. Barker, and U. Amarasinghe, "Water scarcity in the twenty first century," *Water Resources Development*, vol. 15, no. 1, pp. 29 – 42, 1999.

[2] G. Anastasi, M. Conti, and M. D. Francesco, "Extending the lifetime of wireless sensor networks through adaptive sleep," *IEEE Trans. Industrial Informatics*, vol. 5, no. 3, pp. 351–365, 2009.

[3] "Overview of philadelphia's water infrastructure," www.phillywatersheds.org/watershed issues/infrastructure management.

[4] "Control and mitigation of drinking water losses in distribution systems," water.epa.gov/type/drink/pws/smallsystems/upload/Water_Loss_Control_508_FINALDEc.pdf, 2010.

[5] T. Al-Kadi, Z. Al-Tuwaijri, and A. Al-Omran, "Wireless sensor networks for leakage detection in underground pipelines: A survey paper," in *AASNET*, 2013, pp. 491–498.

[6] I. Stoianov, L. Nachman, S. Madden, and T. Tokmouline, "Pipenet: A wireless sensor network for pipeline monitoring," in *IPSN*, 2007, pp. 264–273.

[7] Z. Sun, P. Wang, M. C. Vuran, M. Al-Rodhaan, A. Al-Dhelaan, and I. F. Akyildiz, "MISE-PIPE: magnetic induction-based wireless sensor networks for underground pipeline monitoring," *Ad Hoc Networks*, vol. 9, no. 3, pp. 218–227, 2011.

[8] C. Renner, J. Jessen, and V. Turau, "Lifetime Prediction for Supercapacitor-Powered Wireless Sensor Nodes."

[9] www.bentley.com/en-US/Products/WaterGEMS/Features-List.htm.

[10] D. Moss, J. Hui, and K. Klues, "Low Power Listening, Core Working Group, TEP 105."

[11] M. Khan, M. Iqbal, and J. Quaicoe, "River current energy conversion systems: Progress, prospects and challenges," *Renewable and Sustainable Energy Reviews*, vol. 12, no. 8, pp. 2177–2193, 2008.

[12] www.orbitcoms.com/EverPump.php.

[13] http://ampl.com/products/solvers/.

[14] L. Su, Y. Gao, Y. Yang, and G. Cao, "Towards optimal rate allocation for data aggregation in wireless sensor networks," in *MobiHoc*, 2011.

[15] R. Liu, P. Sinha, and C. E. Koksal, "Joint energy management and resource allocation in rechargeable sensor networks," in *INFOCOM*, 2010, pp. 902–910.

[16] P. R. Bhave and R. Gupta, "Analysis of water distribution networks," 2006.

[17] "Castalia: A Simulator for WSN," castalia.npc.nicta.com.au/.

[18] "Embedded Energy in Water Study 3: End-use Water Demand Profile Final Research Plan," uc-ciee.org/downloads/Eeiswtudy3.pdf, 2009.

[19] "Design Criteria For Water Distribution Systems ," www.wmwd.com/DocumentCenter/Home/View/239, 2011.

[20] J. Kim, G. Sharma, N. Boudriga, and S. S. Iyengar, "SPAMMS: A sensor-based pipeline autonomous monitoring and maintenance system," in *COMSNETS*, 2010, pp. 1–10.

[21] C. Choi and K. Youcef-Toumi, "Robot design for high flow liquid pipe networks," in *IROS*, 2013, pp. 246–251.

[22] A. Kadri, A. Abu-Dayya, D. Trinchero, and R. Stefanelli, "Autonomous sensing for leakage detection in underground water pipelines," in *ICST*, 2011, pp. 639–643.

[23] J. Izquierdo, R. Prez, and P. Iglesias, "Mathematical models and methods in the water industry," *Mathematical and Computer Modelling*, vol. 39, no. 1112, pp. 1353 – 1374, 2004.

[24] Z. Y. Wu, P. Sage, and D. Turtle, "Pressure-dependent leak detection model and its application to a district water system," *Journal of Water Resources Planning and Management*, vol. 136, no. 1, pp. 116 – 128, 2010.

[25] A. Bachir, M. Dohler, T. Watteyne, and K. K. Leung, "MAC essentials for wireless sensor networks," *IEEE Communications Surveys and Tutorials*, vol. 12, no. 2, pp. 222–248, 2010.

[26] C. Tang and C. Raghavendra, "Compression techniques for wireless sensor networks," pp. 207–231, 2004.

[27] A. Pal and A. Nasipuri, "PCOR: A joint power control and routing scheme for rechargeable sensor networks," in *IEEE WCNC*, 2014, pp. 2230–2235.

[28] A. Pal and A. Nasipuri, "DRCS: A distributed routing and channel selection scheme for multi-channel wireless sensor networks," in *IEEE PerSeNS*, 2013, pp. 602–608.

[29] S. Rangwala, R. Gummadi, R. Govindan, and K. Psounis, "Interference-aware fair rate control in wireless sensor networks," in *SIGCOMM*, 2006, pp. 63–74.

[30] C. T. Ee and R. Bajcsy, "Congestion control and fairness for many-to-one routing in sensor networks," in *SenSys*, 2004, pp. 148–161.

[31] E. Shi and A. Perrig, "Designing secure sensor networks," *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 38–43, 2004.

[32] A. Perrig, J. A. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Commun. ACM*, vol. 47, no. 6, pp. 53–57, 2004.