Distributed Quality-Lifetime Maximization in Wireless Video Sensor Networks

Eren Gürses*, Yuan Lin†, Raouf Boutaba*
*School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada
†Center for Quantifiable QoS in Communication Systems, NTNU, N-7491, Trondheim, Norway
{egurses, rboutaba}@uwaterloo.ca, yuan.lin@q2s.ntnu.no

***Abstract*—Owing to the availability of low-cost and low-power CMOS cameras, Wireless Video Sensor Networks (WVSN) has recently become a reality. However video encoding is still a costly process for energy and capacity constrained sensor nodes and this urges the vitality of the control over the network lifetime. In this paper we propose a distributed *quality-lifetime* control algorithm where quality is simply measured by the visual signal quality. In order to formulate the quality-lifetime problem, we consider the Power-Rate-Distortion (P-R-D) model of the video encoder together with the rate control, medium access and routing functions of the underlying communication protocol and formulate it as a Generalized Network Utility Maximization (GNUM) problem. Then we construct the distributed solution based on duality and proximal point methods with necessary convergence analysis. Simulation results support that optimal quality-lifetime control is possible through the distributed algorithm, where the desired point of operation is simply adjusted by the operator via a configuration parameter.**

I. Introduction

Wireless video sensor networks (WVSN) have drawn significant amount of attention in the recent years due to both numerous potential application areas and enhancements they offer to existing WSN applications, such as video surveillance, battlefield awareness, environmental monitoring and industrial process control. [1]. Due to unattended and energy-capacity constrained nature of wireless sensor networks, the design of distributed, energy efficient, self-organizing and optimizing algorithms and communication protocols have become the typical challenges in the recent years. A common factor of all these efforts is to maximize the network lifetime while accomplishing the given task which is generally decided by the measure of utility obtained from the network. In WVSNs a typical utility measure would be the sum of sensor-to-sink Signal-to-Noise-Ratio (SNR) which reflects the visual signal quality and provides a good measure for quality. Hence the quality increases with the amount of data communicated to the sink. However, higher data rate entails higher energy dissipation for capturing, processing and communication in sensor nodes which results in a shorter network lifetime. Hence the design of a quality-lifetime optimization framework for WVSN, necessitate the proper modeling of both video encoding process and the underlying communication mechanisms.

Technical Report CS-2008-19, extended version of a paper submitted to ICC 2009. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

There has been some WVSN prototypes introduced in the recent years such as Cyclops, Panoptes, SensEye [1] however they focus on practical and architectural aspects rather than the distributed system optimization. In [4], authors propose a distributed network lifetime maximization scheme considering video encoder model together with several communication layers. However they use distortion as a constraint not the objective function which may lead to infeasible problems. In [3] authors propose a Generalized Network Utility Maximization (GNUM) based approach to solve the quality-lifetime maximization problem. However their work is not for video sensors and do not employ any Power-Rate-Distortion (P-R-D) model. Furthermore the proposed route selection algorithm continuously oscillates due to the discontinuous mapping from dual variables to primal ones. In this work we consider the P-R-D model of the video encoding process together with the underlying communication protocol functions namely as rate control, medium access and route selection, and formulate it as a quality-lifetime maximization problem based on GNUM. Then we devise a distributed algorithm based on duality and proximal point methods to solve it and provide the thorough convergence analysis which yields the bounds on the selection of step sizes that guarantee the convergence.

In Sec. II we give the system architecture and the GNUM based formulation. In Sec. III we introduce the distributed quality-lifetime maximization algorithm together with the neessary convergence analysis. In Sec. IV we present the simulation results and then concluded in Sec. V.

II. System Architecture

Considering a wireless video sensor network consisting of N nodes and a sink. Let $\mathcal{N}$, $\mathcal{S}$, $\mathcal{L}$ and $\mathcal{R}(s)$ respectively denote the set of nodes, source, links and routes available to source s. Each sensor s generates data at a rate of y_s b/s transmitted over a set of routes $r \in \mathcal{R}(s)$ at rates x_r where $y_s = \sum_{r\in\mathcal{R}(s)} x_r$. Each link $l \in \mathcal{L}$ has a fixed capacity of c_l b/s.

A. Power-Rate-Distortion Model

In a typical WSN node, power is consumed either at the sensor stage for sensing and processing or at the radio stage for communication. In this section we formulate the processing power consumption at the sensor stage through a power-rate-distortion (P-R-D) model [2] where the communication power is later considered in Section II-C. An active video sensor captures and compresses the signal by introducing coding distortion $D_s(\alpha_s, y_s)$ given by (1) as a function of bit rate y_s (in bpp) and I (intra), P (inter) video coding mode rates α_s, $1-\alpha_s$. In the rest of the paper we ignore the additional distortion due to channel losses by assuming the use of some transmit power adaptation technique to minimize receiver bit error rate (BER). In (1), $\delta_s^{(I)}$, $\delta_s^{(P)}$, γ_I, γ_P are given as variance and model accuracy constants for I and P coding modes respectively and $k(\alpha_s, \gamma_I, \gamma_P) \triangleq \frac{\alpha_s}{\gamma_I} + \frac{1-\alpha_s}{\gamma_P}$ is defined.

$$D_s(\alpha_s, y_s) = k(\alpha_s, \gamma_I, \gamma_P)\left[(\delta_s^{(I)}\gamma_I)^{\frac{\alpha_s}{\gamma_I}}(\delta_s^{(P)}\gamma_P)^{\frac{1-\alpha_s}{\gamma_P}} 2^{-2y_s}\right]^{\frac{1}{k(\alpha_s \gamma_I \gamma_P)}} \quad (1)$$

However, in (1) power information is implicit and available through I and P mode encoding power parameters $p_s^{(I)}$ and $p_s^{(P)}$. Then normalizing by $p_s^{(P)}$, we get average normalized encoder power as $p_s = \alpha_s p_s^{(I)}/p_s^{(P)} + (1-\alpha_s)$. After defining $\omega_s = p_s^{(P)}/(p_s^{(P)} - p_s^{(I)})$ and solving for α_s, coding mode rates are found as $\alpha_s = \omega_s(1-p_s)$ once p_s is known. The last step for obtaining the P-R-D model is to simplify (1) by setting $\gamma_I = \gamma_P = 1$ which implies $k(\alpha_s, \gamma_I, \gamma_P) \triangleq 1$. Hence the simplified distortion model is obtained as $D_s(\alpha_s, y_s) = \delta_s^{(P)}\big(\delta_s^{(I)}/\delta_s^{(P)}\big)^{\alpha_s} 2^{-2y_s}$ where we can immediately obtain the P-R-D model in (2) by replacing $\alpha_s = \omega_s(1-p_s)$ for constants κ_s and η_s (3) given for each source and $d_1 = 2\log 2$.

$$D_s(p_s, y_s) = \eta_s \exp(-\kappa_s p_s)\exp(-d_1 y_s) \quad (2)$$

$$\kappa_s = \omega_s \log\left(\frac{\delta_s^{(I)}}{\delta_s^{(P)}}\right),\ \eta_s = \delta_s^{(P)}\left(\frac{\delta_s^{(I)}}{\delta_s^{(P)}}\right)^{\omega_s} \quad (3)$$

B. Joint Medium Access and Congestion Control

In our system model, we don't specify any MAC layer algorithm. Instead we abstract the MAC layer operation through some parameters a_l, ϵ_m and the concept of maximal cliques from graph theory. In that way we integrate the effect of medium access to the proposed distributed network quality-lifetime maximization framework.

We first define an abstract MAC layer parameter $a_l \in [0,1]$ which describes the percent of the physical link capacity c_l used by the MAC layer on any link l for the transmission of the upper layer data In other words, the link can not be used for the transmission of upper layer data during $1-a_l$ percent of the time due to either the capture of medium by another link or the MAC layer overhead such as collisions, control messages, header overhead. Hence the effective MAC layer capacity is $a_l c_l$ and we can now write the *rate constraint* for any link l in the network as $\sum_{r\in\mathcal{R}} h_{l,r}x_r \leq a_l c_l$ for entries $h_{l,r}$ of the routing matrix where $h_{l,r} = 1$ denotes that the link l is part of the route r and 0 otherwise. The constraint can alternatively be written as $\sum_{r\in\mathcal{R}} h_{l,r}x_r/c_l \leq a_l$ for $c_l \geq 0$.

For the medium access part, let M denotes the total number of maximal cliques. Each link is a member of one or more maximal cliques and links in the same clique can not be active at the same time. Hence, using this local conflict information of maximal cliques we construct the global conflict matrix F of dimension $M \times L$ where entries $f_{m,l}$ are equal to 1 if link l is a member of clique m, and 0 otherwise. Finally we define a parameter $\epsilon_m \in [0,1]$ specific to the utilized MAC layer algorithm. It describes the usable percent of the medium within maximal clique m during which links in that clique can use the medium to transmit their (upper layer) data. In other words $1-\epsilon_m$ percent of the time the medium of max clique m is spent to the overhead of the MAC layer algorithm. Hence, in each maximal clique m, the sum of medium access probabilities a_l for conflicting links should satisfy the inequality $\sum_{l\in\mathcal{L}} f_{m,l}a_l \leq \epsilon_m$. Then by replacing a_l with rate constraint, we obtain; *joint rate and medium access constraint* as $\sum_{l\in\mathcal{L}}\sum_{r\in\mathcal{R}} f_{m,l}\big(h_{l,r}x_r/c_l\big) \leq \epsilon_m$.

C. Power-Lifetime Control

In this section, we introduce communication power model and then combine it with the processing power p_s (normalized) in Section II-A to form the power-lifetime control mechanism. For the communication power, we define $\xi_{n,l}$ as the energy consumed per bit on link l of node n where $P_{tx,l}$ and $P_{rx,l}$ are given to be the transmitter and receiver power respectively.

$$\xi_{n,l} = \begin{cases} P_{tx,l}/c_l & \text{, if } l \text{ is an outgoing link of node } n \\ P_{rx,l}/c_l & \text{, if } l \text{ is an incoming link of node } n \\ 0 & \text{, otherwise.} \end{cases} \quad (4)$$

Later for the processing power part, we define the variable $q_{n,s}$ where $q_{n,s} = p_s^{(P)}$ indicates sensor node n is a source sensor s and $q_{n,s} = 0$ otherwise. Note that we allow single video source per node. Then, combining the communication and sensor power, the average power dissipation at node n is $\bar{p}_n = \sum_r\sum_l \xi_{n,l}\, h_{l,r}\, x_r + \sum_{s\in S} q_{n,s}p_s$. Let each node n have an initial energy j_n and the lifetime is given by $T_n = \frac{j_n}{\bar{p}_n}$. We assume any node n runs out of battery results in the failure of the whole network. Hence, we define the network lifetime as the network's operation time until any of the node's energy is depleted and given by $T_{\min} = \min\{T_n | n = 1,\cdots N\}$. Let $v = 1/T_{\min}$ be the inverse network lifetime, then lifetime of each node n satisfies $T_n \geq 1/v$. We define $e_{n,r} = \sum_l \xi_{n,l}h_{l,r}$ in order to simplify the notation where $e_{n,r}$ is the node n's total energy consumption per bit for flow x_r. Then we have the *power constraint* as $\sum_{r\in\mathcal{R}} e_{n,r}x_r + \sum_{s\in S} q_{n,s}p_s \leq j_n v$.

D. Quality-Lifetime Maximization Problem

In this section we introduce the quality-lifetime maximization problem for WVSNs. The objective of the problem has two parts. The first part maximizes the *lifetime* of the network where the second part provides the best total sensor-to-sink signal SNR (quality) as a measure of *quality* for the sensed events. In the first part, network lifetime is maximized by minimizing the inverse lifetime v given in Section II-C through a convex utility function $g(v) = v^2/2\theta$ with some constant θ. For the second part, higher quality is achieved by simply increasing peak signal-to-noise-ratio (PSNR) of the received signal at the sink, which is given by $\text{PSNR} = d2(\log(255^2) - \log(D_s(p_s, y_s)))$ where $d_2 = 10/\log(10)$. Then the overall quality is given by the sum of individual sensor reliabilities $-\sum_s d_2\log(D_s(p_s, y_s))$, where maximizing it is equivalent to minimizing $f(\mathbf{p}, \mathbf{x}) = \sum_s d_2\log(D_s(p_s, y_s))$ for $y_s = \sum_{r\in\mathcal{R}(s)} x_r$. After replacing $D_s(p_s, y_s)$ with (2) and defining $K = d_2\sum_s \log(\eta_s)$, we write the objective function for quality as $f(\mathbf{p},\mathbf{x}) = K - f_1(\mathbf{p}) - f_2(\mathbf{x})$ using functions $f_1(\mathbf{p})$ and $f_2(\mathbf{x})$ in (5).

$$f_1(\mathbf{p}) = d_2 \sum_s \kappa_s p_s \quad ; \quad f_2(\mathbf{x}) = d_1 d_2 \sum_s \sum_{r\in\mathcal{R}(s)} x_r \quad (5)$$

After taking into account the power constraint in Sec. II-C and the joint rate and medium access control constraint in Sec. II-B, the quality-lifetime maximization problem can be expressed as below. While writing the quality-lifetime maximization problem we drop the constant K in the objective function by knowing that the optimal point of the original problem is $\mathcal{P}^* = \hat{\mathcal{P}}^* + K$.

$$\hat{\mathcal{P}}^* = \min_{\mathbf{p},\mathbf{x},v} g(v) - f_1(\mathbf{p}) - f_2(\mathbf{x}) \quad (6)$$

$$\text{s.t.} \sum_{r\in\mathcal{R}} e_{n,r}x_r + \sum_{s\in S} q_{n,s}p_s \le j_n v \ , \forall n \in \mathcal{N} \quad (7)$$

$$\sum_{l\in\mathcal{L}} \sum_{r\in\mathcal{R}} f_{m,l}\big(h_{lr}/c_l\big) x_r \le \epsilon_m \ , \forall m \in \mathcal{M} \quad (8)$$

$$p_{\min} \le p_s \le p_{\max} \ ; \ 0 \le x_r \le x_{\max} \ ; \ v_{\max} \le v \le v_{\min} \quad (9)$$

However dual based methods are not directly applicable, since optimum values for primal variables $\mathbf{p}$ and $\mathbf{x}$ are not immediately available. Hence in the following sections we develop a method to calculate the optimal values of these primal variables which are to be used together with the standard dual based algorithm in order to solve the quality-lifetime maximization problem in a distributive way.

III. Dual-based Distributed Quality-Lifetime Maximization Algorithm

In this section we develop a dual-based distributed algorithm based on the *proximal point* method in order to solve the quality-lifetime maximization problem given in (6)-(9). For that purpose we use the matrix based notation where $\mathbf{H}$, $\mathbf{E}$, $\mathbf{Q}$ and $\mathbf{F}$ are the matrices with entries $h_{l,r}$, $e_{n,r}$, $q_{n,s}$ and $f_{m,l}$ respectively, $\mathbf{C}$=diag$(c_1,\ldots,c_L)$ is the capacity matrix and $\mathbf{x} = [x_1 \ldots x_R]^T$, $\mathbf{p}$=$[p_1 \ldots p_S]^T$, $\boldsymbol{\epsilon}$=$[\epsilon_1 \ldots \epsilon_M]^T$, $\mathbf{j}$=$[j_1 \ldots j_N]^T$, $\boldsymbol{\kappa}$=$[\kappa_1 \ldots \kappa_S]^T$, $\boldsymbol{\eta}$=$[\eta_1 \ldots \eta_S]^T$ are the vectors.

Let, the optimal values of the original problem in (6)-(9) are given by $\mathbf{p}^* = [p_1^* \ldots p_S^*]^T$ and $\mathbf{x}^* = [x_1^* \ldots x_R^*]^T$, then in order to develop the *proximal point* algorithm, we first define new variables $\hat{\mathbf{p}} = [\hat{p}_1 \ldots \hat{p}_S]^T$, $\hat{\mathbf{x}} = [\hat{x}_1 \ldots \hat{x}_R]^T$, then by using them and constants V_{ps}, V_{xr} we obtain the quadratic terms $\frac{1}{2}\sum_s V_{ps}(p_s - \hat{p}_s)^2$ and $\frac{1}{2}\sum_r V_{xr}(x_r - \hat{x}_r)^2$ in order to add to the objective function in (6). These quadratic terms do not effect the optimal point of the original problem since the optimal values for new variables are achieved at $\hat{\mathbf{p}}^*$=$\mathbf{p}^*$ and $\hat{\mathbf{x}}^*$=$\mathbf{x}^*$. Then we represent these quadratic terms in matrix notation by defining diagonal matrices $\mathbf{V}_p$=diag$(V_{p1} \ldots V_{pS})$, $\mathbf{V}_x$=diag$(V_{x1} \ldots V_{xR})$ and using following norm definitions.

$$||\mathbf{p} - \hat{\mathbf{p}}||_{V_p} = (\mathbf{p} - \hat{\mathbf{p}})^T \mathbf{V}_p (\mathbf{p} - \hat{\mathbf{p}}) \quad (10)$$

$$||\mathbf{x} - \hat{\mathbf{x}}||_{V_x} = (\mathbf{x} - \hat{\mathbf{x}})^T \mathbf{V}_x (\mathbf{x} - \hat{\mathbf{x}}) \quad (11)$$

Finally we rewrite quality-lifetime maximization problem as follows which is later shown to be distributively solved by using proximal point algorithm and dual-based methods.

$$\min_{\substack{\mathbf{p},\mathbf{x},v,\\ \hat{\mathbf{p}},\hat{\mathbf{x}}}} g(v) - f_1(\mathbf{p}) - f_2(\mathbf{x}) + \frac{1}{2}||\mathbf{p} - \hat{\mathbf{p}}||_{V_p} + \frac{1}{2}||\mathbf{x} - \hat{\mathbf{x}}||_{V_x} \quad (12)$$

$$\text{s.t. } \mathbf{E}\mathbf{x} + \mathbf{Q}\mathbf{p} \le \mathbf{j}v \quad (13)$$

$$\mathbf{F}\mathbf{C}^{-1}\mathbf{H}\mathbf{x} \le \boldsymbol{\epsilon} \quad (14)$$

$$\mathbf{p}_{\min} \le \mathbf{p} \le \mathbf{p}_{\max} \ ; \ \mathbf{0} \le \mathbf{x} \le \mathbf{x}_{\max} \ ; \ v_{\max} \le v \le v_{\min} \quad (15)$$

We write the Lagrangian as below, after relaxing constraints (13), (14) using $\boldsymbol{\mu}$=$[\mu_1 \ldots \mu_N]^T \in \mathbf{R}_+^N$ and $\boldsymbol{\psi}$=$[\psi_1 \ldots \psi_M]^T \in \mathbf{R}_+^M$ which respectively correspond to power price at nodes and joint congestion-medium access price at max. clique m.

$$\hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu},\boldsymbol{\psi};\hat{\mathbf{p}},\hat{\mathbf{x}})$$

$$= g(v) - f_1(\mathbf{p}) - f_2(\mathbf{x}) + \frac{1}{2}||\mathbf{p} - \hat{\mathbf{p}}||_{V_p} + \frac{1}{2}||\mathbf{x} - \hat{\mathbf{x}}||_{V_x} \quad (16)$$

$$+ \boldsymbol{\mu}^T\mathbf{E}\mathbf{x} + \boldsymbol{\mu}^T\mathbf{Q}\mathbf{p} - \boldsymbol{\mu}^T\mathbf{j}v + \boldsymbol{\psi}^T\mathbf{F}\mathbf{C}^{-1}\mathbf{H}\mathbf{x} - \boldsymbol{\psi}^T\boldsymbol{\epsilon}$$

$$= \sum_s \hat{L}_p(p_s,\boldsymbol{\mu};\hat{p}_s) + \sum_r \hat{L}_x(x_r,\boldsymbol{\mu},\boldsymbol{\psi};\hat{x}_r) + \hat{L}_v(v,\mu) - \boldsymbol{\psi}^T\boldsymbol{\epsilon} \quad (17)$$

In (16), after replacing the functions $f_1(\mathbf{p})$, $f_2(\mathbf{x})$ with (5) and $g(v) = v^2/2\theta$ the Lagrangian becomes separable over $\mathbf{p}$, $\mathbf{x}$ where partial Lagrangians in (17) is given as:

$$\hat{L}_p(p_s,\boldsymbol{\mu};\hat{p}_s) = \sum_n \mu_n q_{ns} p_s - d_2\kappa_s p_s + \frac{V_{ps}}{2}(p_s - \hat{p}_s)^2 \quad (18)$$

$$\hat{L}_x(x_r,\boldsymbol{\mu},\boldsymbol{\psi};\hat{x}_r) = \sum_n \mu_n e_{nr} x_r + \sum_m \sum_l \psi_m f_{ml} \frac{h_{lr}}{c_l} x_r - d_1 d_2 x_r + \frac{V_{xr}}{2}(x_r - \hat{x}_r)^2 \quad (19)$$

$$\hat{L}_v(v,\boldsymbol{\mu}) = v^2/2\theta - \sum_n \mu_n j_n v \quad (20)$$

Following the definition of Lagrangians, dual-based solution in convex optimization is obtained by first finding the dual function $D(\boldsymbol{\mu},\boldsymbol{\psi}) = \min_{\mathbf{p},\mathbf{x},v,\hat{\mathbf{p}},\hat{\mathbf{x}}} \hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu},\boldsymbol{\psi};\hat{\mathbf{p}},\hat{\mathbf{x}})$. Then by maximizing it $\hat{\mathcal{P}}^* = \max_{\boldsymbol{\mu},\boldsymbol{\psi}} D(\boldsymbol{\mu},\boldsymbol{\psi})$ the same primal solution of the problem in (6)-(9) can be obtained.

In Table I distributed solution to the dual-based quality-lifetime maximization problem is given. The algorithm is synchronous and should run once at each time slot t. Steps of the algorithm are designed to run on specific nodes such as sink, source and intermediate sensor nodes. If a node is both a source node and an intermediate sensor node (relay node), then it runs both of the steps in the given order. Note that in steps A.1 and A.3 of the algorithm the primal variables $\mathbf{p}^0(t)$, $\mathbf{x}^0(t)$, $v(t)$ are calculated at sensors and the sink by minimizing partial Lagrangians in (18)-(20) for prices $\boldsymbol{\mu}(t)$ and $\boldsymbol{\psi}(t)$. Following this, in step A.2 the dual updates are done throughout the network by using these optimal primal variable as follows in (21)-(22) in order to maximize the concave dual function. For the dual updates in matrix notation we define dual step size matrices $\mathbf{A}_\mu$ = diag$(\alpha_{\mu 1},\ldots,\alpha_{\mu N})$ and $\mathbf{A}_\psi$ = diag$(\alpha_{\psi 1},\ldots,\alpha_{\psi M})$.

$$\boldsymbol{\mu}(t+1) = \big[\boldsymbol{\mu}(t) + \mathbf{A}_\mu(\mathbf{E}\mathbf{x}^0(t) + \mathbf{Q}\mathbf{p}^0(t) - \mathbf{j}v(t))\big]^+ \quad (21)$$

$$\boldsymbol{\psi}(t+1) = \big[\boldsymbol{\psi}(t) + \mathbf{A}_\psi(\mathbf{F}\mathbf{C}^{-1}\mathbf{H}\mathbf{x}^0(t) - \boldsymbol{\epsilon})\big]^+ \quad (22)$$

Finally in step A.4 of the algorithm, each source uses the updated prices $\boldsymbol{\mu}(t+1)$ and $\boldsymbol{\psi}(t+1)$ and recalculates its

TABLE I
DUAL-BASED DISTRIBUTED QUALITY-LIFETIME MAXIMIZATION

Initialize: Set $t=0$ and initialize $v(t)$, $\mu_n(t)$, $\psi_m(t)$, $\hat{p}_s(t)$ and $\hat{x}_r(t)$.
A.1. Sensor Algorithm (Source):
Sensors encode and transmit according to variables x_r^0, $\alpha_s{=}\omega_s(1{-}p_s^0)$

$$p_s^0(t) = \arg\min_{p_s} \hat{L}_p(p_s, \mu(t); \hat{p}_s(t)) \text{ using (18)}$$
$$x_r^0(t) = \arg\min_{x_r} \hat{L}_x(x_r, \mu(t), \psi(t); \hat{x}_r(t)) \text{ using (19)}$$

A.2. Sensor Algorithm (Intermediate):
Each intermediate sensor node n and maximal clique m that node n is involved in, update prices as follows using (21) and (22) and feed back this information to relevant sources and the sink

$$\mu_n(t{+}1){=}\Big[\mu_n(t){+}\alpha_{\mu n}(\sum_r e_{nr}x_r^0(t) + \sum_s q_{ns}p_s^0(t){-}j_n v(t))\Big]^+$$
$$\psi_m(t{+}1){=}\Big[\psi_m(t){+}\alpha_{\psi m}(\sum_l\sum_r f_{ml}\frac{h_{lr}}{c_l}x_r^0(t){-}\epsilon_m)\Big]^+$$

A.3. Sink Algorithm:
Calculates $v(t+1) = \arg\min_v \hat{L}_v(v, \mu(t+1))$ using (20) and broadcast it to sensors
A.4. Sensor Algorithm (Source):
Each source recalculates p_s, x_r for new values of $\boldsymbol{\mu}(t+1)$, $\boldsymbol{\psi}(t+1)$

$$p_s^1(t) = \arg\min_{p_s} \hat{L}_p(p_s, \mu(t+1); \hat{p}_s(t)) \text{ using (18)}$$
$$x_r^1(t) = \arg\min_{x_r} \hat{L}_x(x_r, \mu(t+1), \psi(t+1); \hat{x}_r(t)) \text{ using (19)}$$

then maximizes the concave function of $\hat{p}$, $\hat{x}$ returned by above min. as:

$$\hat{p}_s(t+1) = \hat{p}_s(t) + \beta_{\hat{p}s}(p_s^1(t) - \hat{p}_s(t)) \text{ using (23)}$$
$$\hat{x}_r(t+1) = \hat{x}_r(t) + \beta_{\hat{x}r}(x_r^1(t) - \hat{x}_r(t)) \text{ using (24)}$$

Set time $t=t+1$ and continue with step A.1

optimal primal variable $\mathbf{p}^1(t)$, $\mathbf{x}^1(t)$ that minimize partial Lagrangians (18)-(19). Using the updated primal variables, the primal variables of the proximal point algorithm $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$ are updated as follows with primal step size matrices $\mathbf{B}_{\hat{p}} = \text{diag}(\beta_{\hat{p}1}, \ldots, \beta_{\hat{p}S})$ and $\mathbf{B}_{\hat{x}} = \text{diag}(\beta_{\hat{x}1}, \ldots, \beta_{\hat{x}R})$.

$$\hat{\mathbf{p}}(t+1) = \hat{\mathbf{p}}(t) + \mathbf{B}_{\hat{p}}(\mathbf{p}^1(t) - \hat{\mathbf{p}}(t)) \quad (23)$$
$$\hat{\mathbf{x}}(t+1) = \hat{\mathbf{x}}(t) + \mathbf{B}_{\hat{x}}(\mathbf{x}^1(t) - \hat{\mathbf{x}}(t)) \quad (24)$$

However without the convergence analysis, it is not possible to know for which values of the step sizes the algorithm in Table I converges. Therefore in the rest of this section, we conduct the convergence analysis to discover the bounds on step sizes that lead to global optimum. Hence, first we provide auxilary Lemmas 1-5 and finally give these bounds in Prop. 6. Due to lack of space we only provide the proof of Prop. 6 and skip the proofs of others which are more straightforward. For the rest of the paper we define following minimizers. $[\mathbf{p}^0(t),\mathbf{x}^0(t),v(t)]{=}\text{argmin}_{\mathbf{p},\mathbf{x},v}\hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu}(t),\boldsymbol{\psi}(t);\hat{\mathbf{p}}(t),\hat{\mathbf{x}}(t))$, $[\mathbf{p}^1(t),\mathbf{x}^1(t),v(t{+}1)]{=}\text{argmin}_{\mathbf{p},\mathbf{x},v}\hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu}(t{+}1),\boldsymbol{\psi}(t{+}1);$ $\hat{\mathbf{p}}(t),\hat{\mathbf{x}}(t))$.

Lemma 1 *For fixed $\hat{\mathbf{p}}$, $\hat{\mathbf{x}}$, following inequalities hold for any minimizer $[\mathbf{p}_1,\mathbf{x}_1,v_1]$ and $[\mathbf{p}_2,\mathbf{x}_2,v_2]$ of $\hat{L}(.)$, respectively for the prices $\boldsymbol{\mu}_1$, $\boldsymbol{\psi}_1$ and $\boldsymbol{\mu}_2$, $\boldsymbol{\psi}_2$ (i.e. given that $[\mathbf{p}_1,\mathbf{x}_1,v_1] = \arg\min_{\mathbf{p},\mathbf{x},v}\hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu}_1,\boldsymbol{\psi}_1;\hat{\mathbf{p}},\hat{\mathbf{x}})$ and $[\mathbf{p}_2,\mathbf{x}_2,v_2] = \arg\min_{\mathbf{p},\mathbf{x},v}\hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu}_2,\boldsymbol{\psi}_2;\hat{\mathbf{p}},\hat{\mathbf{x}})$)*

1) $||\mathbf{p}_2{-}\mathbf{p}_1||_{\mathbf{V}_p} \le \big|\big|\mathbf{Q}^T(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)\big|\big|_{\mathbf{V}_p^{-1}}$
2) $||\mathbf{x}_2{-}\mathbf{x}_1||_{\mathbf{V}_x} \le \big|\big|\mathbf{E}^T(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)+\mathbf{FC}^{-1}\mathbf{H}(\boldsymbol{\psi}_2-\boldsymbol{\psi}_1)\big|\big|_{\mathbf{V}_x^{-1}}$

Proof: We give the proof of part 2 where (25) gives the difference of gradients (i.e. $\nabla_x = [\partial/\partial x_1, \ldots, \partial/\partial x_R]^T$) of $\hat{L}(.)$ evaluated at minimizers $[\mathbf{p}_2,\mathbf{x}_2,v_2]$ and $[\mathbf{p}_1,\mathbf{x}_1,v_1]$.

$$[\nabla_x f_2(\mathbf{x}_2) - \nabla_x f_2(\mathbf{x}_1)] - \mathbf{V}_x(\mathbf{x}_2-\mathbf{x}_1) = \mathbf{E}^T(\boldsymbol{\mu}_2-\boldsymbol{\mu}_1)+\mathbf{FC}^{-1}\mathbf{H}(\boldsymbol{\psi}_2-\boldsymbol{\psi}_1) \quad (25)$$

We multiply (25) with $\mathbf{V}_x^{-1}$ from left and multiply the result with the transpose of (25) from left again. In the resulting expression, we note that $||\nabla_x f_2(\mathbf{x}_2)-\nabla_x f_2(\mathbf{x}_1)||_{\mathbf{V}_x^{-1}} \ge 0$ and $-2[\nabla_x f_2(\mathbf{x}_2){-}\nabla_x f_2(\mathbf{x}_1)]^T(\mathbf{x}_2{-}\mathbf{x}_1) \ge 0$ due to property of norm and concavity of $f_2(\mathbf{x})$ respectively. Hence part 2 of Lemma 1 holds. Similarly part 1 is shown by using ∇_p. ■

Lemma 2 *The optimal point $[\mathbf{p}^*,\mathbf{x}^*,v^*]$ of (6)-(9) for $\boldsymbol{\mu}^*$, $\boldsymbol{\psi}^*$ and the minimizer $[\mathbf{p}^1(t),\mathbf{x}^1(t),v(t+1)]$, satisfy following :*

1) $[\nabla_p f_1(\mathbf{p}^1(t)){-}\nabla_p f_1(\mathbf{p}^*)]{-}\mathbf{V}_p(\mathbf{p}^1(t){-}\hat{\mathbf{p}}(t)){=}\mathbf{Q}^T(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)$
2) $[\nabla_x f_2(\mathbf{x}^1(t)) - \nabla_x f_2(\mathbf{x}^*)] - \mathbf{V}_x(\mathbf{x}^1(t)-\hat{\mathbf{x}}(t))$
$= \mathbf{E}^T(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*){+}\mathbf{FC}^{-1}\mathbf{H}(\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}^*)$
3) $[g'(v(t+1)) - g'(v^*) = (\boldsymbol{\mu}(t+1)-\boldsymbol{\mu}^*)^T\mathbf{j}$

Proof: The optimal point $[\mathbf{p}^*,\mathbf{x}^*,v^*]$ with optimal prices $\boldsymbol{\mu}^*$, $\boldsymbol{\psi}^*$ is also the optimizer of (12)-(15) where $\hat{\mathbf{p}}^* = \mathbf{p}^*$, $\hat{\mathbf{x}}^* = \mathbf{x}^*$. By taking the gradient ∇_p of $\hat{L}(.)$, evaluating it at the minimizer $[\mathbf{p}(t),\mathbf{x}(t),v(t)]$ and at the optimal point $[\mathbf{p}^*,\mathbf{x}^*,v^*]$, and then subtracting them we obtain part 1. Similary part 2 is obtained by using ∇_x. ■

Lemma 3 *Assume $\boldsymbol{\mu}(t)$, $\boldsymbol{\psi}(t)$ and $\boldsymbol{\mu}(t+1)$, $\boldsymbol{\psi}(t+1)$ are prices updated according to (21)-(22) and $[\mathbf{p}^0(t),\mathbf{x}^0(t),v(t)]$ is the minimizer, then the following relations hold:*

1) $(\boldsymbol{\mu}(t+1)-\boldsymbol{\mu}^*)^T\big[\mathbf{E}\mathbf{x}^0(t)+\mathbf{Q}\mathbf{p}^0(t)-\mathbf{j}v(t)\big]$
$\le(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\big[\mathbf{E}(\mathbf{x}^0(t){-}\hat{\mathbf{x}}^*){+}\mathbf{Q}(\mathbf{p}^0(t){-}\hat{\mathbf{p}}^*){-}\mathbf{j}(v(t){-}v^*)\big]$
2) $(\boldsymbol{\psi}(t+1)-\boldsymbol{\psi}^*)^T\big[\mathbf{FC}^{-1}\mathbf{H}\mathbf{x}^0(t)-\boldsymbol{\epsilon}\big]$
$\le(\boldsymbol{\psi}(t+1)-\boldsymbol{\psi}^*)^T\big[\mathbf{FC}^{-1}\mathbf{H}(\mathbf{x}^0(t)-\hat{\mathbf{x}}^*)\big]$

Proof: We keep terms on the right hand side of part 1 which are equivalent to the ones on the left, and investigate the $(\boldsymbol{\mu}^*{-}\boldsymbol{\mu}(t{+}1))^T(\mathbf{E}\hat{\mathbf{x}}^*{+}\mathbf{Q}\hat{\mathbf{p}}^*{-}\mathbf{j}v^*)$. We can replace $\hat{\mathbf{p}}^* = \mathbf{p}^*$ and $\hat{\mathbf{x}}^* = \mathbf{x}^*$ since they are equal at optimality. The equality ${\boldsymbol{\mu}^*}^T(\mathbf{E}\mathbf{x}^*+\mathbf{Q}\mathbf{p}^*-\mathbf{j}v^*) = 0$ holds due to complementary slackness. On the other hand, since the dual function is concave and maximized over $\boldsymbol{\mu}$ and $\boldsymbol{\psi}$ with maximum values at $\boldsymbol{\mu}^*$ and $\boldsymbol{\psi}^*$, then for any value of $\boldsymbol{\mu}(t{+}1) \ne \boldsymbol{\mu}^*$ the product $\boldsymbol{\mu}(t+1)^T(\mathbf{E}\mathbf{x}^*+\mathbf{Q}\mathbf{p}^*-\mathbf{j}v^*) \le 0$ is true and it follows the inequality in part 1. Similarly part 2 can be shown. ■

Lemma 4 *Given that $[\mathbf{p}^1(t),\mathbf{x}^1(t),v(t{+}1)]$ is the minimizer of $\hat{L}(\mathbf{p},\mathbf{x},v,\boldsymbol{\mu}(t{+}1),\boldsymbol{\psi}(t{+}1);\hat{\mathbf{p}}(t),\hat{\mathbf{x}}(t))$ following relations hold*

1) $||\hat{\mathbf{p}}(t+1)-\hat{\mathbf{p}}^*||_{\mathbf{B}_p^{-1}\mathbf{V}_p} - ||\hat{\mathbf{p}}(t)-\hat{\mathbf{p}}^*||_{\mathbf{B}_p^{-1}\mathbf{V}_x u}$
$\le \big|\big|\mathbf{p}^1(t)-\hat{\mathbf{p}}^*\big|\big|_{\mathbf{V}_p} - ||\hat{\mathbf{p}}(t)-\hat{\mathbf{p}}^*||_{\mathbf{V}_p}$
2) $||\hat{\mathbf{x}}(t+1)-\hat{\mathbf{x}}^*||_{\mathbf{B}_x^{-1}\mathbf{V}_x} - ||\hat{\mathbf{x}}(t)-\hat{\mathbf{x}}^*||_{\mathbf{B}_x^{-1}\mathbf{V}_x}$
$\le \big|\big|\mathbf{x}^1(t)-\hat{\mathbf{x}}^*\big|\big|_{\mathbf{V}_x} - ||\hat{\mathbf{x}}(t)-\hat{\mathbf{x}}^*||_{\mathbf{V}_x}$

Proof: We can rewrite the update of $\hat{\mathbf{p}}$ in step A.4 of Table I as $\hat{p}_s(t+1) = (1-\beta_{\hat{p}s})\hat{p}_s(t) + \beta_{\hat{p}s}p_s^1(t)$. Then we subtract p_s^* from both sides and use triangle inequality. Since

$\beta_{\hat{p}s}$ and $(1-\beta_{\hat{p}s}) \in [0,1]$ their squares are smaller than themselves and the terms that are multiplied with them are all positive, then the following inequality holds.

$$(\hat{p}_s(t{+}1){-}p_s^*)^2 \leq (1{-}\beta_{\hat{p}s})(\hat{p}_s(t){-}p_s^*)^2{+}\beta_{\hat{p}s}(p_s^1(t){-}p_s^*)^2$$

Multiplying both sides by V_{ps}, the following inequality holds

$$\frac{V_{ps}}{\beta_{\hat{p}s}}(\hat{p}_s(t{+}1){-}p_s^*)^2 - \frac{V_{ps}}{\beta_{\hat{p}s}}(\hat{p}_s(t){-}p_s^*)^2 \leq V_{ps}(p_s^1(t){-}p_s^*)^2 - V_{ps}(\hat{p}_s(t){-}p_s^*)^2 \quad (26)$$

since $V_{ps}, \beta_{\hat{p}s} \geq 0$ which concludes part 1 of Lemma 4. ■

Lemma 5 *Projection mappings for constrained variables $\boldsymbol{\mu}(t)$, $\boldsymbol{\psi}(t)$ given by (21)-(22) satisy the following relations:*

1) $(\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}^*)^T \mathbf{A}_\mu^{-1}(\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}(t)) \leq (\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}^*)^T(\mathbf{E}\mathbf{x}^0(t) + \mathbf{Q}\mathbf{p}^0(t) - \mathbf{j}v(t))$
2) $(\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}^*)^T \mathbf{A}_\psi^{-1}(\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}(t)) \leq (\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}^*)^T(\mathbf{F}\mathbf{C}^{-1}\mathbf{H}\mathbf{x}^0(t) - \boldsymbol{\epsilon})$

Proof: Given a convex set $\mathcal{X}$, for any point $y \in \mathcal{X}$ and $x \notin \mathcal{X}$ the orthogonal projection mapping $[x]^+$ onto the convex set $\mathcal{X}$ satisfies $([x]^+ - y)([x]^+ - x) \leq 0$. Then knowing that $\boldsymbol{\mu}^*$ is in the convex set, following inequality holds for (21)

$$(\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}^*)^T \mathbf{A}_\mu^{-1} . \Big[\boldsymbol{\mu}(t{+}1){-}\big(\boldsymbol{\mu}(t){+}\mathbf{A}_\mu(\mathbf{E}\mathbf{x}^0(t){+}\mathbf{Q}\mathbf{p}^0(t){-}\mathbf{j}v(t))\big)\Big] \leq 0 \quad (27)$$

and follows part 1 of Lemma 5. Part 2 is proven similarly. ■

Note that $\sum_m f_{ml}$ and $\sum_r h_{lr}$ respectively refer to the number of max. cliques for link l and the number of routes using link l. Then we define, $L_\mathcal{R}{=}\max_r \sum_l (h_{lr}/c_l)\sum_m f_{ml}$, $L_\mathcal{M}{=}\max_m \sum_l (f_{ml}/c_l)\sum_r h_{lr}$, $J_T{=}\sum_n j_n$, $j'{=}\max_n j_n$, $E_\mathcal{R}{=}\max_r \sum_n e_{nr}$, $E_\mathcal{N}{=}\max_n \sum_r e_{nr}$, $Q_\mathcal{S}{=}\max_s \sum_n q_{ns}$ and $Q_\mathcal{N}{=}\max_n \sum_s q_{ns}$.

Proposition 6 *The dual-based distributed quality-lifetime maximization algorithm (i.e. Table I) converges to a stationary point for the following conditions on step sizes $\alpha_{\psi m}$, $\alpha_{\mu n}$ and the proximal point algorithm parameters V_{xr}, V_{ps}.*

$$\min_m \left[\frac{1}{\alpha_{\psi m}}\right] \geq 2L_\mathcal{R}L_\mathcal{M}\max_r\left[\frac{1}{V_{xr}}\right] \quad (28)$$

$$\min_m \left[\frac{1}{\alpha_{\mu m}}\right] \geq Q_\mathcal{S}Q_\mathcal{N}\max_s\left[\frac{1}{V_{ps}}\right] + 2E_\mathcal{R}E_\mathcal{N}\max_r\left[\frac{1}{V_{xr}}\right] + \theta J_T j' \quad (29)$$

Proof: First in (30) we define a Lyapunov function $U(t)$, then to prove Prop. 6 we show that $U(t{+}1){-}U(t){\leq}0, \forall t$.

$$U(t) = ||\boldsymbol{\mu}(t) - \boldsymbol{\mu}^*||_{\mathbf{A}_\mu^{-1}} + ||\boldsymbol{\psi}(t) - \boldsymbol{\psi}^*||_{\mathbf{A}_\psi^{-1}} + ||\hat{\mathbf{p}}(t) - \hat{\mathbf{p}}^*||_{\mathbf{B}_p^{-1}\mathbf{V}_p} + ||\hat{\mathbf{x}}(t) - \hat{\mathbf{x}}^*||_{\mathbf{B}_x^{-1}\mathbf{V}_x} + \frac{1}{\theta}(v(t) - v^*)^2 \quad (30)$$

We first write (31) from the generalization of law of cosines to some normed vector space with norm $||.||_{\mathbf{A}_\mu^{-1}}$, then by applying part 1 of Lemma 5 and Lemma 3 we obtain (32):

$$\begin{aligned} &||\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}^*||_{\mathbf{A}_\mu^{-1}} - ||\boldsymbol{\mu}(t) - \boldsymbol{\mu}^*||_{\mathbf{A}_\mu^{-1}} \quad (31)\\ &= -||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}(t)||_{\mathbf{A}_\mu^{-1}} {+} 2(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{A}_\mu^{-1}(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}(t))\\ &\leq -||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}(t)||_{\mathbf{A}_\mu^{-1}} {+} 2\Big[(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{E}(\mathbf{x}^0(t){-}\hat{\mathbf{x}}^*) \quad (32)\\ &\quad +(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{Q}(\mathbf{p}^0(t){-}\hat{\mathbf{p}}^*) {-} (\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{j}(v(t){-}v^*)\Big] \end{aligned}$$

Similarly we can write the relation for $\boldsymbol{\psi}$ after replacing part 2 of the Lemma 5 and Lemma 3.

$$\begin{aligned} &||\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}^*||_{\mathbf{A}_\psi^{-1}} - ||\boldsymbol{\psi}(t){-}\boldsymbol{\psi}^*||_{\mathbf{A}_\psi^{-1}}\\ &= -||\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}(t)||_{\mathbf{A}_\psi^{-1}} {+} 2(\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}^*)^T\mathbf{A}_\psi^{-1}(\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}(t))\\ &\leq -||\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}(t)||_{\mathbf{A}_\mu^{-1}} {+} 2(\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}^*)^T(\mathbf{F}\mathbf{C}^{-1}\mathbf{H}(\mathbf{x}^0(t){-}\hat{\mathbf{x}}^*)) \quad (33) \end{aligned}$$

We replace the terms $(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{E}$ and $(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{Q}$ in (32) with Lemma 2 parts 2 and 1 respectively. Then for $g(v) = v^2/2\theta$ part 3 of Lemma 2 becomes $(v(t) - v^*) = (\boldsymbol{\mu}(t) - \boldsymbol{\mu}^*)^T\mathbf{j}\theta$ and add the resulting espression to (33). After canceling the similar terms, we obtain the following relation. Note that for the following expressions we define $\mathbf{J}_\theta \triangleq \mathbf{j}\theta\mathbf{j}^T$

$$\begin{aligned} &||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*||_{\mathbf{A}_\mu^{-1}} {-} ||\boldsymbol{\mu}(t){-}\boldsymbol{\mu}^*||_{\mathbf{A}_\mu^{-1}} {+} ||\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}^*||_{\mathbf{A}_\psi^{-1}} {-} ||\boldsymbol{\psi}(t){-}\boldsymbol{\psi}^*||_{\mathbf{A}_\psi^{-1}}\\ &\leq -||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}(t)||_{\mathbf{A}_\mu^{-1}} - ||\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}(t)||_{\mathbf{A}_\psi^{-1}}\\ &+2\Big([\nabla_p f_1(\mathbf{p}^1(t)){-}\nabla_p f_1(\mathbf{p}^*)]^T(\mathbf{p}^0(t){-}\hat{\mathbf{p}}^*) {-} (\mathbf{p}^1(t){-}\hat{\mathbf{p}}(t))^T\mathbf{V}_p(\mathbf{p}^0(t){-}\hat{\mathbf{p}}^*)\\ &+[\nabla_x f_2(\mathbf{x}^1(t)){-}\nabla_x f_2(\mathbf{x}^*)]^T(\mathbf{x}^0(t){-}\hat{\mathbf{x}}^*) {-} (\mathbf{x}^1(t){-}\hat{\mathbf{x}}(t))^T\mathbf{V}_x(\mathbf{x}^0(t){-}\hat{\mathbf{x}}^*)\\ &-(\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*)^T\mathbf{J}_\theta(\boldsymbol{\mu}(t){-}\boldsymbol{\mu}^*)\Big) \quad (34) \end{aligned}$$

Then we can rewrite the terms related to v in $U(t{+}1){-}U(t)$ by using $(v(t) - v^*) = (\boldsymbol{\mu}(t) - \boldsymbol{\mu}^*)^T\mathbf{j}\theta$ as given above.

$$\frac{1}{\theta}\Big[(v(t{+}1){-}v^*)^2 {-} (v(t){-}v^*)^2\Big] = ||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}^*||_{\mathbf{J}_\theta} {-} ||\boldsymbol{\mu}(t){-}\boldsymbol{\mu}^*||_{\mathbf{J}_\theta} \quad (35)$$

Finally the terms related to $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$ in $U(t+1) - U(t)$ are directly given in Lemma 4. Before continuing with the final step, the following equation holds for any set of vectors x, y, w, z from a normed vector space $\mathcal{V}$ with norm $||.||_A$. $||x{-}w||_A {-} ||y{-}w||_A = ||x{-}z||_A {-} ||y{-}z||_A {-} 2(x{-}y)^T A(w{-}z)$. Using this property and constructing $U(t{+}1){-}U(t)$ by using (34), (35) and Lemma 5, we obtain the following.

$$\begin{aligned} &U(t{+}1){-}U(t) \leq -||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}(t)||_{\mathbf{A}_\mu^{-1}} - ||\boldsymbol{\psi}(t{+}1){-}\boldsymbol{\psi}(t)||_{\mathbf{A}_\psi^{-1}}\\ &+\big|\big|\mathbf{p}^1(t){-}\mathbf{p}^0(t)\big|\big|_{\mathbf{V}_p} {-} \big|\big|\hat{\mathbf{p}}(t){-}\mathbf{p}^0(t)\big|\big|_{\mathbf{V}_p} {+} \big|\big|\mathbf{x}^1(t){-}\mathbf{x}^0(t)\big|\big|_{\mathbf{V}_x} {-} \big|\big|\hat{\mathbf{x}}(t){-}\mathbf{x}^0(t)\big|\big|_{\mathbf{V}_x}\\ &+||\boldsymbol{\mu}(t{+}1){-}\boldsymbol{\mu}(t)||_{\mathbf{J}_\theta} {-} 2||\boldsymbol{\mu}(t){-}\boldsymbol{\mu}^*||_{\mathbf{J}_\theta} {+} 2[\nabla_p f_1(\mathbf{p}^1(t)){-}\nabla_p f_1(\mathbf{p}^*)]^T(\mathbf{p}^0(t){-}\hat{\mathbf{p}}^*)\\ &+2[\nabla_x f_2(\mathbf{x}^1(t)){-}\nabla_x f_2(\mathbf{x}^*)]^T(\mathbf{x}^0(t){-}\hat{\mathbf{x}}^*) \quad (36) \end{aligned}$$

In order to prove the convergence in Prop. 6 we need to show $U(t{+}1){-}U(t)$ in (36) is negative. The terms $-\big|\big|\hat{\mathbf{p}}(t){-}\mathbf{p}^0(t)\big|\big|_{\mathbf{V}_p}$, $-\big|\big|\hat{\mathbf{x}}(t){-}\mathbf{x}^0(t)\big|\big|_{\mathbf{V}_x}$ and $-2||\boldsymbol{\mu}(t){-}\boldsymbol{\mu}^*||_{\mathbf{J}_\theta}$ are all negative and conform to the convergence condition. Next, we replace the terms $\big|\big|\mathbf{p}^1(t){-}\mathbf{p}^0(t)\big|\big|_{\mathbf{V}_p}$ and $\big|\big|\mathbf{x}^1(t){-}\mathbf{x}^0(t)\big|\big|_{\mathbf{V}_x}$ with values in Lemma 1. Finally, due to the special form of functions $f_1(\mathbf{p})$ and $f_2(\mathbf{x})$, the gradient terms in (36) are

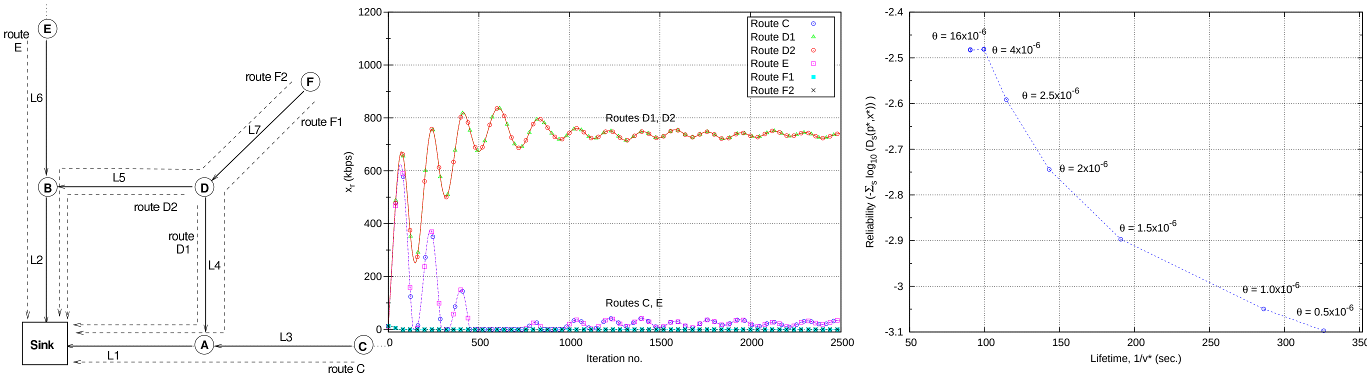

Fig. 1. (a) Topology, (b) Source rates on each route, (c) Optimal quality-lifetime control

$[\nabla_p f_1(\mathbf{p}^1(t)) - \nabla_p f_1(\mathbf{p}^*)] = 0$ and $[\nabla_x f_2(\mathbf{x}^1(t)) - \nabla_x f_2(\mathbf{x}^*)] = 0$. Then we relax the condition on term at the last line of (36):

$$2[\nabla_x f_2(\mathbf{x}^1(t)) - \nabla_x f_2(\mathbf{x}^*)]^T(\mathbf{x}^0(t) - \hat{\mathbf{x}}^*) \leq \left|\left|\mathbf{E}^T(\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}(t)) - \mathbf{F}\mathbf{C}^{-1}\mathbf{H}(\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}(t))\right|\right|_{\mathbf{V}_x^{-1}} \quad (37)$$

By using these properties, we obtain the convergence condition $U(t+1) - U(t) \leq 0$ for (36) as follows:

$$(\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}(t))^T\big[2\mathbf{E}\mathbf{V}_x^{-1}\mathbf{E}^T + \mathbf{Q}\mathbf{V}_p^{-1}\mathbf{Q}^T + \mathbf{J}_\theta - \mathbf{A}_\mu^{-1}\big](\boldsymbol{\mu}(t+1) - \boldsymbol{\mu}(t)) + (\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}(t))^T\left[2\mathbf{F}\mathbf{C}^{-1}\mathbf{H}\mathbf{V}_x^{-1}(\mathbf{F}\mathbf{C}^{-1}\mathbf{H})^T - \mathbf{A}_\psi^{-1}\right](\boldsymbol{\psi}(t+1) - \boldsymbol{\psi}(t)) \leq 0 \quad (38)$$

In order to satisfy the convergence condition in (38) we investigate the bounds on step sizes $\mathbf{V}_p^{-1}$, $\mathbf{V}_x^{-1}$ and $\mathbf{A}_\mu^{-1}$, $\mathbf{A}_\psi^{-1}$ which guarantees the terms in square brackets to be negative semi-definite. For convenience, we define $\mathbf{G} \triangleq \mathbf{F}\mathbf{C}^{-1}\mathbf{H}$ then for any vector $\mathbf{z} = [z_1 \dots z_M]^T$ we need to show the second term in (38) to be $2\mathbf{z}^T\mathbf{G}\mathbf{V}_x^{-1}\mathbf{G}^T\mathbf{z} \leq \mathbf{z}^T\mathbf{A}_\psi^{-1}\mathbf{z}$. We first write the right hand side as $\mathbf{z}^T\mathbf{A}_\psi^{-1}\mathbf{z} = \sum_m (1/\alpha_{\psi m}) z_m^2$. For the left hand side after defining $L_\mathcal{R} = \max_r[\sum_m G_{mr}]$ and $L_\mathcal{M} = \max_m[\sum_r G_{mr}]$ where $G_{mr} = \sum_l f_{ml} h_{lr}/c_l$.

$$\sum_r \frac{2}{V_{xr}}\Big[\sum_m G_{mr} z_m\Big]^2 \leq \sum_r \frac{2}{V_{xr}}\Big(\sum_m G_{mr}\Big)\sum_m G_{mr} z_m^2$$
$$\leq 2L_\mathcal{R} \max_r\left[\frac{1}{V_{xr}}\right]\sum_m\Big(\sum_r G_{mr}\Big)z_m^2 \leq 2L_\mathcal{R}L_\mathcal{M}\max_r\left[\frac{1}{V_{xr}}\right]\sum_m z_m^2$$

Then we obtain $\min_m\big[\frac{1}{\alpha_{\psi m}}\big] \geq 2L_\mathcal{R}L_\mathcal{M}\max_r\big[\frac{1}{V_{xr}}\big]$ in Prop. 6 to make the second term in (38) negative semi-definite.

Similarly for any vector $\hat{\mathbf{z}} = [\hat{z}_1 \dots \hat{z}_N]^T$ we need to show the first term in (38) to be $\hat{\mathbf{z}}^T\left[2\mathbf{E}\mathbf{V}_x^{-1}\mathbf{E}^T + \mathbf{Q}\mathbf{V}_p^{-1}\mathbf{Q}^T + \mathbf{J}_\theta\right]\hat{\mathbf{z}} \leq \hat{\mathbf{z}}^T\mathbf{A}_\mu^{-1}\hat{\mathbf{z}}$ Note that maximum one video source per node is allowed. Then after defining $E_\mathcal{R} = \max_r\left[\sum_n e_{nr}\right]$, $E_\mathcal{N} = \max_n\left[\sum_r e_{nr}\right]$, $Q_\mathcal{S} = \max_s\left[\sum_n q_{ns}\right]$, $Q_\mathcal{N} = \max_n\left[\sum_s q_{ns}\right]$, $J_T = \sum_n j_n$ and $j' = \max_n j_n$, the following bound on step sizes is obtained.

$$\min_m\left[\frac{1}{\alpha_{\mu m}}\right] \geq Q_\mathcal{S}Q_\mathcal{N}\max_s\left[\frac{1}{V_{ps}}\right] + 2E_\mathcal{R}E_\mathcal{N}\max_r\left[\frac{1}{V_{xr}}\right] + \theta J_T j'$$

■

IV. Simulation Results

In order to illustrate the results of the proposed quality-lifetime maximization algorithm we use the simple topology given in Fig 1(a). Nevertheless it can easily scale up to larger networks since the algorithm is distributed. In the given network, we have $N{=}6$ nodes, $L{=}7$ links, $S{=}4$ sources $\mathcal{S}{=}\{C, D, E, F\}$ and $R{=}6$ routes $\mathcal{R}{=}\{C, D1, D2, , E, F1, F2\}$. We assume a simple MAC scheme which only prevents simultaneous use of transmitter and/or receiver of a node on different links. Hence there are $M{=}4$ maximal cliques with $\mathcal{M}{=}\{\{L1, L2\}, \{L1, L3, L4\}, \{L2, L5, L6\}, \{L4, L5, L7\}\}$. We set the distances between adjacent nodes to 50 m. Receiver power of all links $L1, \dots, L7$ are fixed to 1.2 mW. where the transmitter powers are set as $P_{tx} = [3.0\ 3.0\ 4.5\ 2.5\ 2.5\ 4.5\ 3.0]^T$ mW. according to some transmit power adaptation policy. We assume a constant noise level of 8 dB on all links. Hence considering the interference, noise, transmit power and distance we calculate link capacities approximately as $\mathbf{c}{=}[1.87\ 1.87\ 1.20\ 1.73\ 1.73\ 1.20\ 0.04]^T$ Mbps. Finally we assign initial energy of $j = [2\ 2\ 2\ 2\ 2\ 2]^T$ joules for each node where video encoder power is set as $p_s^{(I)}{=}4$ mW. and $p_s^{(I)}{=}20$ mW. We assume all video encoders are encoding at QCIF resolution at 5 frames/s which are necessary to convert the data rate to bpp (bits per pixel) from bps. Finally $\beta_\psi{=}\beta_\mu{=}1$, $V_p{=}4$, $V_x{=}64$ and unless otherwise stated $\theta{=}2\times10^{-6}$. Given the above values, we may decide the entities $E_\mathcal{R}{=}0.0137$, $E_\mathcal{N}{=}0.019$, $L_\mathcal{R}{=}3.4419$, $L_\mathcal{M}{=}6.6124$, $Q_\mathcal{N}{=}Q_\mathcal{S}{=}0.02$, $J_T{=}12$ and $j'{=}2$ which are used to find the bounds $\alpha_\psi \leq 1.406$ and $\alpha_\mu \leq 6.4\times10^3$ from Prop. 6.

In all experiments we use $\mathcal{S}{=}\{C, D, E, F\}$ and allow these nodes to capture, compress and transmit video data. In our first experiment, we observe from Fig. 1(b) that in the optimal solution most of the video data is requested from source D through routes D1 and D2 with 740 kbps on each. Then a small portion is requested from sensors C and E with 34 kbps on each. This result is expected since D1 and D2 are the least costly routes which is observed by comparing individual link costs found by dividing transmit power P_{tx} to the link capacities $\mathbf{c}$, On the other hand we can observe the power control at the video encoder from Fig. 2 by only following the curves for $j_D = 2$ joules. As expected, at optimality source D requires a lower energy level for encoder (i.e. higher

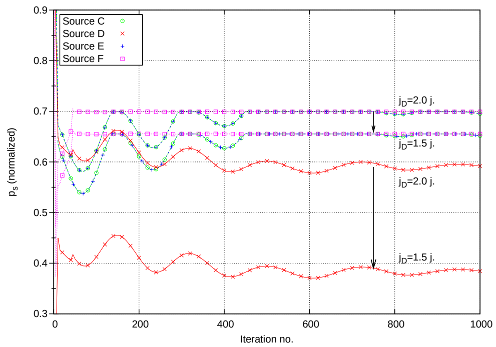

Fig. 2. Optimal power adaptation for video encoder where $\alpha_s = \omega_s(1-p_s)$

rate of I-coded blocks) than the sources C, D, since it spends comparably much higher energy for the communication. I-coding rate for the video encoder of source D is directly calculated by $\alpha_s = \omega_s(1-p_s)$ and it shows I-coding rate increases with the decrease in the budget for p_s.

In the second experiment we run the quality-lifetime maximization algorithm for various values of θ=[0.5 1.0 1.5 2.0 2.5 4.0 16.0]$\times 10^{-6}$ which is simply used to find different trade-off policies between network lifetime and quality as illustrated in Fig. 1(c). As the operator asks for a longer lifetime, by setting a lower θ, the network could optimally sacrifice from the quality to meet the longer lifetime requirement (or vice versa). Finally in the last simulation, we repeat the first experiment for the same parameters except the initial energy $j_D = 1.5$ j. of source D. In this case in Fig. 2 the encoder power of D is dramatically reduced, as compared to the small decrease in the encoder power of others sources that compensates the additional traffic assigned to sensors C, E.

V. Conclusion

In this work we propose a distributed optimal quality-lifetime control algorithm for wireless video sensor networks. The problem is modeled as a GNUM problem by taking into consideration the video encoder parameters, source rates and routes, and the channel contention. Results show that distributed optimal quality-lifetime control could be achieved for different policies described by the θ.

References

[1] I. F. Akyildiz, et.al. "A survey on wireless multimedia sensor networks," *Computer Networks*, vol. 51, no. 4, pp.921-960, Mar. 2007.

[2] Y. Lin, et. al. "Optimal joint power-rate adaptation for error resilient video coding," *IS&T/SPIE VCIP*, Jan. 2008.

[3] H. Nama, et.al. "Optimal utility-lifetime trade-off in wireless sensor networks: Characterization and distributed algorithms," *IEEE ICC*, 2006.

[4] Y. F. He, et.al. "Network lifetime maximization in wireless visual sensor networks using a distributed algorithm," *Proc. of IEEE ICME* , 2007.