Optimizing Networked Situational Awareness

Jean Abou Rahal[*], Gustavo de Veciana[*], Takayuki Shimizu[**], and Hongsheng Lu[**]

[*]Dept. of Electrical and Computer Engineering, The University of Texas at Austin, TX
[**]TOYOTA InfoTechnology Center, U.S.A., Inc.
Email: jeanabourahal@utexas.edu, gustavo@ece.utexas.edu, tshimizu@us.toyota-itc.com and hlu@us.toyota-itc.com

***Abstract*—This paper proposes a framework to explore the optimization of applications where a distributed set of nodes/sensors, e.g., automated vehicles, collaboratively exchange information over a network to achieve real-time situational-awareness. To that end we propose a reasonable proxy for the usefulness of possibly delayed sensor updates and their sensitivity to the network resources devoted to such exchanges. This enables us to study the joint optimization of (1) the application-level update rates, i.e., how often and when sensors update other nodes, and (2), the transmission resources allocated to, and resulting delays associated with, exchanging updates. We first consider a network scenario where nodes share a single resource, e.g., an ad hoc wireless setting where a cluster of nodes, e.g., platoon of vehicles, share information by broadcasting on a single collision domain. In this setting we provide an explicit solution characterizing the interplay between network congestion and situational awareness amongst heterogeneous nodes. We then extend this to a setting where such clusters can also exchange information via a base station. In this setting we characterize the optimal solution and develop a natural distributed algorithm based on exchanging congestion prices associated with sensor nodes' update rates and associated network transmission rates. Preliminary numerical evaluation provides initial insights on the trade-offs associated with optimizing situational awareness and the proposed algorithm's convergence.**

I. Introduction

In this paper we study the fundamental characteristics of systems aimed at achieving real-time situational-awareness based on distributed sensing resources. In particular we explore the joint selection of sensor update rates/policies and the allocation of communication resources towards optimizing Networked Situational-Awareness (NSA).

As an example of such a system, we consider automated vehicles leveraging collaborative sensing. Each car has access to on-board sensing resources, e.g., mmwave radar, LIDAR, cameras which provide a local perspective on their dynamic environment. Unfortunately such sensing modalities are tied to the availability of Line-of-Sight (LOS) views, meaning that certain key regions may be obstructed, e.g., a vehicle may not be able to see what is in front of the vehicle ahead of it, or a vehicle may wish to have redundant points of view of its environment to provide improved tracking and/or detection reliability. To overcome this challenge one might consider enabling vehicles to engage in collaborative sensing, wherein they exchange raw (or processed) sensor data with each other towards improving each vehicle's situational-awareness [1]. Such an approach would potentially involve sharing substantial amounts of information amongst nearby vehicles, possibly overloading communication resources. Network congestion or other transmission/processing delays in turn reduce the timeliness of the shared information, compromising the ability of automated vehicles to make reliable real-time decisions. Indeed the sensitivity of collaborative sensing systems to both the latency and capacity of the underlying communication network has motivated the industry to develop 5G wireless standards for Ultra-Reliable Low Latency Communications (URLLC).

The challenges of achieving real-time situational awareness through collaborative sensing in a communication constrained setting are many and involve several fundamental questions, including:

1) How often and when should sensing nodes update their neighbors regarding their respective environments?
2) What is an appropriate metric (or proxy thereof) to quantify situational-awareness and help drive the fair allocation of resources?
3) How should network resources be allocated among competing nodes' updates so as to optimize the overall nodes' situational awareness?

In order to study such systems, we require a well-defined metric. As discussed in more detail below, the Age-of-Information (AoI) has emerged as a simple intuitive metric: it measures how old relative to current time is the most recently received sensor update. This is, of course, only *loosely* tied to situational-awareness. Other more traditional metrics are tied to the achievable distortion/error, e.g. the Mean Square Error (MSE) of an estimated sensor node's "state" at a remote node. As we will see, these metrics are roughly aligned and provide the starting point of this paper.

Related Work. There has been substantial interest in modeling systems involving the timely monitoring of remote processes over a network. The novelty of our work lies in the study of optimizing networked situational-awareness.
Age-of-Information (AoI), as discussed in [2], was introduced in the early 2010s as a measure which quantifies the freshness of the information a node has about a remote node's state. This metric became popular because it better represents the information freshness versus traditional latency/delay. Techniques to quantify and minimize AoI, or simply *Age* have been extensively studied in previous work, see examples, [3], [4], [5], [6] and [7]. In particular [4] studies how to optimally manage the freshness of information updates sent from a *single* source node to a destination, via a channel.

The recent work of [8] focuses on what is perhaps a more natural metric for tracking scenarios. The setting considered involves a single node monitoring a process (Brownian motion) and sending updates over a network (single queue) to a remote node which creates its best estimate for the process based on the received updates. The paper poses and solves the problem of determining an optimal update strategy subject to a constraint on the long term rate of updates, where the cost is given by the time average MSE of the remote site's estimate for the process. Although this is an extremely simplified model, it gives a fundamental characterization of the problem at hand, and will serve to motivate our networked problem formulation.

The general approach proposed in this paper is based on ideas underlying resource allocation in today's communication networks. Specifically work connecting the allocations achieved by transport protocols such as TCP to utility maximization, see e.g., [9], [10] and [11] for an in depth survey. However, our paper differs from this body of work in that it addresses the joint optimization of sensor nodes' update policies and network resource allocation. As we shall see, the setting involves congestion constraints that are not easily decomposable but capture the underlying character of the problem at hand.

Contributions of this paper. In this paper we propose a framework to explore the optimization of networked situational awareness. We study the joint optimization of both the application-level update rate, i.e., how often and when sensors update other nodes, and the transmission resources allocated to, and resulting delays associated with, sharing nodes' updates. We first consider a network scenario where nodes share a single resource, e.g., an ad hoc wireless setting where a cluster of nodes, e.g., platoon of vehicles, share information by broadcasting on a single collision domain, and find closed form expressions for both the update and transmission rates associated with this scenario. We then extend this to a setting where such clusters can in addition exchange information via a base station. In this setting we characterize the optimal solution and develop a natural distributed algorithm based on exchanging congestion prices associated with sensor nodes' update rates and associated network transmission rates. We conclude with a set of preliminary numerical evaluations to explore the algorithm's convergence and character of the resources' allocations.

Organization. The paper is organized as follows. In Section II we motivate and propose an appropriate utility function for situational awareness. Section III describes our system model for a cluster of nodes broadcasting updates to each other over a shared ad-hoc wireless network. Section IV expands our model to include clusters of nodes which can further communicate through network infrastructure. In Section V, we design a dual decomposition algorithm used to jointly optimize sensor nodes' update rates and network transmission rates. Section VI provides preliminary numerical results and analysis, and finally Section VII concludes the paper.

II. Modeling Networked Situational-Awareness

In this section we develop a simplified model for real-time situational awareness in a collaborative sensing system. We focus on a setting where sensing nodes are monitoring "independent" processes and updating their peers accordingly.

As a starting point we consider the AoI metric in a simple idealized setting. Suppose a sensor node periodically generates updates every $1/f$ seconds and each one is delayed by exactly d seconds before reaching the remote node. The time-varying AoI at the remote node is shown in Fig.1.

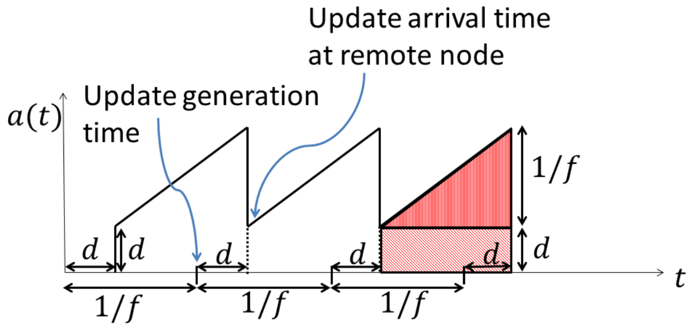

Fig. 1. Transition from AoI to situational awareness metric

This model is idealized in that (1) updates are generated periodically while in practice they could have been generated opportunistically, e.g. based on the degree of change in the underlying process, and (2) network delays are assumed to be fixed, and (3) the focus is on AoI being the appropriate metric. The time average AoI for this idealized process is given by

$$\text{AoI} = \frac{1}{2f} + d. \tag{1}$$

To address these limitations, let us consider the stylized result in [8]. The setting is as follows: a sensing node monitors and samples from a Brownian Motion $(W_t, t \geq 0)$ with variance σ^2. This nodes updates another remote node of the observed processes' changing state, which it does by transmitting an update over a communication link. The updates are known to have i.i.d. service times $(Y_i, i = 1, 2, \ldots)$ with the same distribution as a random variable Y. Further the sensor node is aware of the state of the link, i.e., busy or not. The key result developed in [8] is a characterization of an optimal update strategy, i.e., one that minimizes the time average MSE at the remote node subject to a constraint f on the long term frequency of updates. The optimal updating policy is parameterized by a parameter θ and can be described as follows. The optimal policy generates updates at times $(S_i,\ i = 1, 2 \ldots)$, given by

$$S_{i+1} = \inf\{t \geq S_i + Y_i : |W_t - W_{S_i}| \geq \sqrt{\theta}\}, \tag{2}$$

i.e., the policy waits until the previous update was successfully transmitted, and then sends an update once the change in the process exceeds $\sqrt{\theta}$. The optimal θ is characterized by the following theorem.

Theorem 1: [8] For a given constraint on the update rate f and distribution for the i.i.d. packet service times Y, the optimal θ is given by the solution to the following equation

$$\mathbb{E}[\max(\theta, W_Y^2)] = \max\left[\frac{\sigma^2}{f}, \frac{\mathbb{E}[\max(\theta^2, W_Y^4)]}{2\theta}\right], \tag{3}$$

where W_Y corresponds to the distribution of a Brownian Motion $(W_t, t \geq 0)$ sampled at a random time Y. The optimal MSE is then given by

$$\text{MSE}_{\text{opt}} = \frac{\mathbb{E}[\max(\theta^2, W_Y^4)]}{6\mathbb{E}[\max(\theta, W_Y^2)]} + \sigma^2\mathbb{E}[Y]. \tag{4}$$

This explicit elegant characterization of an optimal updating policy captures both the role of variability of the observed process as well as the impact of packet delays. Also underlying this result is the basic observation that the sensor node should never generate an update when the channel is busy, as the update would simply wait in the queue for transmission. The limitations of this result should also be clear. In particular only a single sensor node is considered with a dedicated transmission link, along with a possibly artificial constraint on the long term rate of updates f the node can generate.

Suppose that the update service times/delays are fixed to d, then one can show after some somewhat intricate approximations (left out due to space constraints) that the optimal threshold θ and associated MSE in Theorem 1 are roughly

$$\theta \approx \frac{\sigma^2}{f} \text{ and } \mathrm{MSE_{opt}} \approx \sigma^2\big(\frac{1}{6f}+d\big). \tag{5}$$

Note that this update threshold θ matches the optimal sampling threshold derived in [12] for the case where $d=0$. These approximate results make explicit the role played by various system parameters. As can be seen, the achievable MSE is lower bounded by $\sigma^2 d$, i.e., no matter how high the sensor update rate f is, it can not overcome the MSE arising due to the delay (service time) d to communicate with the remote node. This brings into focus the critical role that latency plays in networks supporting real-time situational awareness. Still as the allowable update rate f increases the optimal updating policy can make the MSE close to this lower bound. Indeed the reduction in MSE is inversely proportional to f.
Note that for the AoI model discussed earlier, if the observed process were a Brownian Motion, the MSE is equal to $\sigma^2 \times AoI$. Following from Eq.(1), we have that the MSE= $\sigma^2\big(\frac{1}{2f}+d\big)$. Thus the advantage of having an opportunistic update policy as exhibited in Eq.(2), versus a periodic updating policy with the same frequency of updates is the change in the factor multiplying $1/f$ from $\frac{1}{2}$ to $\frac{1}{6}$.

Motivated by the above results, we propose the following parametric model for situational-awareness. Suppose that a sensor update has a size ν bits, and suppose that the sensor has a dedicated link with capacity r bps. Then the delay d to transmit an update is given by $d=\frac{\nu}{r}$. Also note that if the update rate is f (Hz), then the associated average bit rate ρ (bps) generated by the sensor node is given by $\rho = f\nu$. With these new variables the approximate optimal MSE given in Eq.(5) is

$$\mathrm{MSE_{opt}} = \big(\frac{\sigma^2\nu}{6}\big)\frac{1}{\rho} + (\sigma^2\nu)\frac{1}{r},$$

with a similar functional result for AoI. This captures the impact of allowing an increased update rate ρ and/or provisioning a link with an increased capacity r on the MSE / AoI. Note that $\rho \le r$.

Based on these observations we propose the following proxy metric that captures the situational-awareness cost.

Definition. Given an update data rate ρ (bps) and link transmission rate r (bps), we model the Situational Awareness Error (SAE) of a node, $s(.,.)$ as follows:

$$s(\rho, r) = \frac{a}{\rho} + \frac{b}{r}, \tag{6}$$

where $b \ge a > 0$ are constants.
Remark. Through the parameters a and b, this model can capture the salient characteristics of the underlying system. For example, the variability of the underlying process (captured by σ^2) that a sensor node is monitoring would scale a, b. Also, different types of sensor nodes, e.g. video/imaging, LIDAR, might generate updates of different sizes, which would also scale a, b. Finally, the relative values of a, b model the nature of the update policy being used, e.g., deterministic, opportunistic, or other.

The figure below exhibits the "on/off" dynamics of the dedicated transmission link. A simple Corollary to the result developed to prove Theorem 1 gives the characteristics of this process.

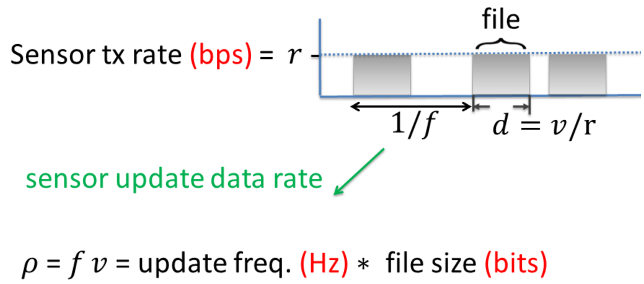

Fig. 2. System model: sensor update process and SAE model

Corollary 1: Under the optimal policy given in both Theorem 1 and Eq.(2) for an update rate constraint f and deterministic packet service times $d = \nu/r$, the stationary dynamics of the communication link correspond to an "on/off" alternating renewal process which has an average "on/off" cycle time of $1/f$ and "on" period of length d during which the link transmits at rate r. Whence the fraction of time the link is busy is $\rho/r < 1$.

In the next section, we shall leverage this basic result to study a more general setting.

III. NSA Optimization on a Shared Broadcast Network

As a first step, we consider a cluster of sensor nodes $\mathcal{N}$ sharing a single broadcast resource (single collision domain). Each node broadcasts updates to all the other nodes. Since different nodes within the cluster are located in different positions, the broadcast rate, μ_n of each node $n \in \mathcal{N}$ may be different, e.g., a node more centrally located within the cluster might have a higher broadcast rate .

Below we consider the problem of jointly optimizing the sensor nodes' update rates $\boldsymbol{\rho} = (\rho_n : n \in \mathcal{N})$ where $\rho_n = f_n \nu_n$ and transmission rates $\boldsymbol{r} = (r_n : n \in \mathcal{N})$. To that end, we define an appropriate cost function along with appropriate capacity constraints.

Objective function. The situational awareness error SAE of node n is modeled as

$$s_n(\rho_n, r_n) = \frac{a_n}{\rho_n} + \frac{b_n}{r_n},$$

where $b_n \ge a_n > 0$, $\forall n \in \mathcal{N}$. The network's overall SAE is then given by

$$g(\boldsymbol{\rho}, \boldsymbol{r}) = \sum_{n\in\mathcal{N}} s_n(\rho_n, r_n) = \sum_{n\in\mathcal{N}} \frac{a_n}{\rho_n} + \frac{b_n}{r_n}.$$

As should be clear, increasing a sensor node's update rate ρ_n and/or transmission rate r_n decreases the SAE as seen at the nodes that it is updating, leading to an improved network situational awareness. Further, note that the overall SAE is convex encoding a degree of fairness across the SAE's of the cluster's nodes.

Congestion constraints. As seen in Corollary 1, if nodes operate on dedicated links, each will act as a stationary alternating renewal process. At any random time, sensor node n could be "on" with probability ρ_n/r_n and transmitting at rate r_n. Otherwise it is "off". We model the state of a sensor node n at a typical time via a Bernoulli random variable $X_n \sim$ Bernoulli (ρ_n/r_n), where ρ_n/r_n represents the fraction of time the link is busy sending node n's update, as described in Corollary 1. We shall further make the following assumption.

Assumption 1: (Independence of sensor nodes' processes). We shall assume that the sensor nodes' states are independent, e.g., the underlying processes they observe are independent.

In our model, if at some point in time a sensor node is transmitting at rate r_n, it will require a fraction of the shared broadcast resource, r_n/μ_n, and if all the nodes were active, to ensure all the nodes' transmissions can be supported, we require that

$$\sum_{n\in\mathcal{N}} \frac{r_n}{\mu_n} \leq 1.$$

However since not all the nodes are active all the time, we will impose a relaxed chance constraint [13] which ensures that with high probability, $1-\epsilon$ (for ϵ very small), the sensor nodes' update transmissions can be supported. In particular, we require

$$P\left(\sum_{n\in\mathcal{N}} X_n \frac{r_n}{\mu_n} > 1\right) < \epsilon. \tag{7}$$

Using the Hoeffding bound, one can show (see Appendix IX-A) that a sufficient condition for the above constraint to be satisfied, is given in the following Lemma.

Lemma 1: Under Assumption 1, the network congestion constraint Eq.(7) is satisfied if

$$\sum_{n\in\mathcal{N}} \frac{\rho_n}{\mu_n} + \omega \parallel \mathbf{r} \parallel_{\boldsymbol{\mu},2} \leq 1, \tag{8}$$

where $\omega = \sqrt{-\frac{1}{2}\ln(\epsilon)}$ and $\parallel \mathbf{x} \parallel_{\boldsymbol{\mu},q} := \left(\sum_{n\in\mathcal{N}} \left(\frac{x_n}{\mu_n}\right)^q\right)^{\frac{1}{q}}$ which for $q \geq 1$ is a weighted, by positive reciprocals of $\boldsymbol{\mu}^{-1}$, norm.

Given Lemma 1, one can formulate the following optimization problem of NSA, and solve for both the update and transmission rates.

Problem 1: (**NSA optimization - single shared resource**)

$$\min_{\boldsymbol{\rho},\boldsymbol{r}} \{g(\boldsymbol{\rho},\boldsymbol{r}) \mid \sum_{n\in\mathcal{N}} \frac{\rho_n}{\mu_n} + \omega \parallel \mathbf{r} \parallel_{\boldsymbol{\mu},2} \leq 1, \quad \boldsymbol{\rho} \leq \boldsymbol{r} \leq \boldsymbol{\mu}\}. \tag{9}$$

Proposition 1: The NSA optimization Problem 1 is convex with a unique solution, which for ϵ small enough, e.g., $e^{-72} \leq \epsilon \leq e^{-2}$, is given by:

$$\rho_n^* = \frac{\alpha_n}{\parallel \boldsymbol{\alpha} \parallel_1 + \sqrt{\omega} \parallel \boldsymbol{\beta} \parallel_{\frac{4}{3}}} \mu_n, \quad r_n^* = \frac{1}{\sqrt{\omega}} \frac{\beta_n}{\alpha_n} \left(\frac{\parallel \boldsymbol{\beta} \parallel_{\frac{4}{3}}}{\beta_n}\right)^{\frac{1}{3}} \rho_n^*,$$

for all $n \in \mathcal{N}$, and where $\boldsymbol{\alpha} = (\alpha_n = \sqrt{\frac{a_n}{\mu_n}} : n \in \mathcal{N})$ and $\boldsymbol{\beta} = (\beta_n = \sqrt{\frac{b_n}{\mu_n}} : n \in \mathcal{N})$.

Proposition 1 is proven in Appendix IX-B.

The relatively simple closed form given above, is obtained by relaxing the constraint $\boldsymbol{\rho} \leq \boldsymbol{r} \leq \boldsymbol{\mu}$ and verifying that under the congestion constraint and the assumption that both $b_n \geq a_n > 0$, $\forall n \in \mathcal{N}$, and $0 < \omega \leq 1$, it will be satisfied.

Remark. To get further insight on the NSA problem, consider the homogeneous case, where all the nodes share the same a_n, b_n parameters and broadcast capacity μ. In this case the optimal sensor node update and transmission rates are given by

$$\rho^* = \left(\frac{1}{1+\sqrt{\omega}\sqrt{\frac{b}{a}}N^{-\frac{1}{4}}}\right)\frac{\mu}{N} \quad \text{and} \quad r^* = \sqrt{\frac{b}{a}}\frac{1}{\sqrt{\omega}}N^{\frac{1}{4}}\rho^*,$$

where $N = |\mathcal{N}|$.

As can be seen, for fixed μ, as N grows, the update rate ρ^* behaves as $\frac{\mu}{N}$ while the optimal transmission rate r^* behaves as $\frac{1}{N^{\frac{3}{4}}}$. Intuitively, we might argue that as the number of sensing nodes in the network increases, optimizing NSA requires that the probability of each node staying "on" decreases as $\frac{1}{N^{\frac{1}{4}}}$, while the transmission rate allocated to each user experiences a less stringent decrease, i.e., as $\frac{1}{N^{\frac{3}{4}}}$, i.e., each user still transmits at a high rate to reduce update delays.

Another interesting setting is one where the broadcast capacity scales in N, i.e., $\mu = \kappa N$, where $\kappa > 0$ is a constant. In that case, ρ^* converges to κ, while r^* increases on the order of $N^{\frac{1}{4}}$. Intuitively, when the broadcast capacity scales linearly, each node can increase its transmission rate, which reduces both its probability of being 'on' and the update delay.

IV. NSA Optimization for Infrastructure Assisted Inter-Cluster Communication

In this section, we extend our previous model to include multiple clusters which can further exchange updates via a base station. The setting is exhibited in Fig. 3.

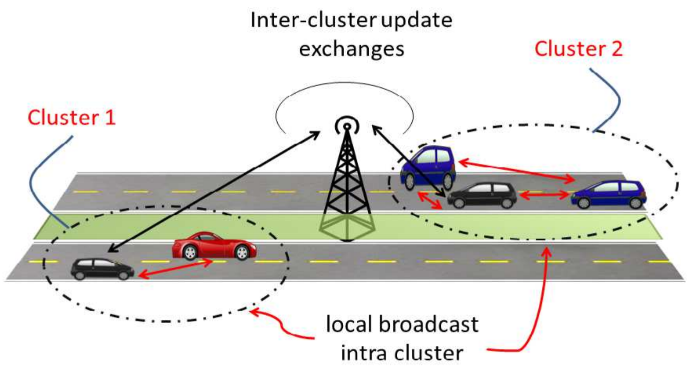

Fig. 3. Infrastructure assisted intra-cluster update exchanges.

In particular the sensor nodes distribute their updates as follows: (1) *Intra-Cluster broadcast*: sensor nodes within the same cluster communicate among themselves via a local broadcast. (2) *Inter-Cluster broadcast*: clusters can share updates with each other by transmitting up to the BS which in turn can broadcast them down to other clusters. We will assume that inter-cluster and intra-cluster communications do not interfere with each other, i.e., operate on orthogonal frequency bands, and that clustering is such that inter-cluster broadcasts do not interfere with each other.

Let $\mathcal{C}$ denote a set of clusters, and $\mathcal{N}_c$ be the set of nodes in cluster $c \in \mathcal{C}$. We provide further discussion regarding the network under consideration.

- Broadcasting is inherently unreliable, thus the intra-cluster and base station inter-cluster broadcast could provide an additional level of reliability for intra-cluster update sharing.

- The uplink transmissions to the base station could be performed in various ways. One possibility is that each sensing node directly sends the update to the base station. Another one is that each cluster selects a *cluster head* which is in charge of forwarding the updates generated by any node in the cluster to the base station. The cluster head might be selected to have either a good connectivity to the nodes in the cluster and/or a good uplink capacity to the BS, thus reducing congestion on the BS uplink.
- For base station downlink transmissions, one could consider either an omni-directional broadcast, where all the nodes in the network receive the update forwarded from the BS instantaneously, or, one could assume that the BS uses a directional type of broadcasting, where it directs its broadcast to a single cluster, one cluster at a time. This might be necessary to ensure better reliability and higher transmission rates from the BS to the nodes. Doing so would require the base station to use several transmissions on the downlink (in fact, one for each cluster).

While the above options can be handled in our framework, below we proceed with a simple and straightforward version: each node $n \in \mathcal{N}_c$ shares its update to all the other nodes in the network. The nodes within the same cluster receive the update via both intra-cluster and base station level broadcast. The nodes in other clusters receive the update via the BS broadcast alone. Each node sends its updates to the BS which then broadcasts them to all the nodes in $\mathcal{N}$. We also make the following assumption:

Assumption 2: (*Cut-through* assumption). We assume that BS uplink/downlink relaying of an update incurs no relaying delay. Below we consider the problem of jointly optimizing the sensor nodes' update rates $\boldsymbol{\rho} = (\rho_n : n \in \mathcal{N})$ and transmission rates $\boldsymbol{r} = (r_n : n \in \mathcal{N})$. We shall also define $\boldsymbol{\rho}^c = (\rho_n : n \in \mathcal{N}_c)$ and $\boldsymbol{r}^c = (r_n : n \in \mathcal{N}_c)$.

The overall cost function SAE, $g(\boldsymbol{\rho}, \boldsymbol{r})$, is the same as that defined earlier.

We denote the intra-cluster broadcast rate of a sensor node n to the nodes in its cluster by μ_n^a, and the uplink/dowlink peak rates from/to particular nodes by μ_n^u and μ_n^d, respectively. We let $\boldsymbol{\mu^a} = (\mu_n^a : n \in \mathcal{N})$. Similarly $\boldsymbol{\mu^u} = (\mu_n^u : n \in \mathcal{N})$, and $\boldsymbol{\mu^d} = (\mu_n^d : n \in \mathcal{N})$. We assume that μ_n^d is the same for all the nodes and hence equal to μ^d.

Congestion constraints. As seen in Corollary 1, if nodes operate in isolation, each will act as a stationary alternating renewal process.

In our infrastructure assisted network model, if at some point in time, node $n \in \mathcal{N}_c$ is transmitting an update at rate r_n, three main constraints need to be satisfied. The first one is dictated by each cluster's resources: each node will require a fraction of its cluster's resources, given by r_n/μ_n^a, and to ensure all the nodes' transmissions can be supported, we require that

$$\sum_{n \in \mathcal{N}_c} \frac{r_n}{\mu_n^a} \leq 1.$$

The second and third constraints are set to avoid congestion at the base station, i.e. the activity at the base station (which is receiving/broadcasting updates) must be supported for all nodes $n \in \mathcal{N}$ on both the uplink and the downlink, and must satisfy

$$\sum_{n \in \mathcal{N}} \frac{r_n}{\mu_n^u} \leq 1 \text{ and } \sum_{n \in \mathcal{N}} \frac{r_n}{\mu^d} \leq 1.$$

However, since not all nodes are active at the same time, we shall, once again, impose a chance constraint [13] which ensures with high probability that the sensor node's update transmissions can be supported. In particular, we require that,

$$P\Big(\sum_{n \in \mathcal{N}_c} X_n \frac{r_n}{\mu_n^a} > 1\Big) < \epsilon, \ \forall c \in \mathcal{C},$$

$$P\Big(\sum_{n \in \mathcal{N}} X_n \frac{r_n}{\mu_n^u} > 1\Big) < \epsilon \text{ and } P\Big(\sum_{n \in \mathcal{N}} X_n \frac{r_n}{\mu^d} > 1\Big) < \epsilon,$$

where $X_n \sim$ Bernoulli (ρ_n/r_n) for all $n \in \mathcal{N}$. As shown in Lemma 1, sufficient constraints for these to be satisfied can be found based on the following norms:

$$\sqrt{\sum_{n \in \mathcal{N}_c} \big(\frac{r_n}{\mu_n^a}\big)^2} = \| \mathbf{r}^c \|_{\boldsymbol{\mu^a},2},$$

$$\sqrt{\sum_{n \in \mathcal{N}} \big(\frac{r_n}{\mu_n^u}\big)^2} = \| \mathbf{r} \|_{\boldsymbol{\mu^u},2} \quad \text{and} \quad \sqrt{\sum_{n \in \mathcal{N}} \big(\frac{r_n}{\mu^d}\big)^2} = \| \mathbf{r} \|_{\boldsymbol{\mu^d},2} .$$

The joint NSA optimization problem with infrastructure assistance can be formulated as follows.

Problem 2: (**NSA optimization for infrastructure assisted setting**).

$$\begin{aligned}\min_{\boldsymbol{\rho},\boldsymbol{r}}\{g(\boldsymbol{\rho}, \boldsymbol{r}) \quad | \quad & \sum_{n \in \mathcal{N}_c} \frac{\rho_n}{\mu_n} + \omega \| \mathbf{r}^c \|_{\boldsymbol{\mu^a},2} \leq 1, \forall c \in \mathcal{C},\\ & \sum_{n \in \mathcal{N}} \frac{\rho_n}{\mu_n^u} + \omega \| \mathbf{r} \|_{\boldsymbol{\mu^u},2} \leq 1,\\ & \sum_{n \in \mathcal{N}} \frac{\rho_n}{\mu^d} + \omega \| \mathbf{r} \|_{\boldsymbol{\mu^d},2} \leq 1\}.\end{aligned}$$

Proposition 2: Under Assumption 1, Problem 2 is convex with a unique solution characterized by first order optimality conditions which gives the following solution: For all $n \in \mathcal{N}$,

$$\rho_n^* = \sqrt{\frac{a_n}{\frac{\lambda_c^a}{\mu_n^a} + \frac{\lambda_b^u}{\mu_n^u} + \frac{\lambda_b^d}{\mu^d}}}, \tag{10}$$

$$r_n^* = \sqrt[3]{\frac{b_n}{\omega\big(\frac{\lambda_c^a}{(\mu_n^a)^2 \|\mathbf{r}^c\|_{\boldsymbol{\mu^a},2}} + \frac{\lambda_b^u}{(\mu_n^u)^2 \|\mathbf{r}\|_{\boldsymbol{\mu^u},2}} + \frac{\lambda_b^d}{(\mu^d)^2 \|\mathbf{r}\|_{\boldsymbol{\mu^d},2}}\big)}}, \tag{11}$$

where $\lambda_c^a, \ c \in \mathcal{C}$, are Lagrange multipliers associated with the intra-cluster congestion constraint and λ_b^u and λ_b^d are associated with BS uplink and downlink congestion constraints, respectively.

Remark. An interesting observation is that while ρ_n^* depends only on congestion prices, r_n^* depends also on other nodes' transmission rates.

In the next section, we propose a distributed algorithm to determine the optimal joint sensor node update and transmission rates' allocation across the sensor nodes.

V. NSA Algorithm

The algorithm works as follows. Each cluster of nodes $c \in \mathcal{C}$ updates its Lagrange multiplier λ_c^a which we refer to as *cluster price*, while the base station updates the Lagrange multipliers, λ_b^u and λ_b^d corresponding to *uplink/downlink prices*, respectively. Meanwhile each sensor node $n \in \mathcal{N}$ responds by updating its sensor update and transmission rates ρ_n and r_n, respectively.

Suppose that each cluster elects a single node to serve as the cluster head. Its main role will be to compute the *cluster price* and establish a direct connection to exchange price information with the base station. λ_c^a is updated at the cluster head, while λ_b^u and λ_b^d are updated at the base station. λ_b^u and λ_b^d are shared with the cluster heads, who forward λ_c^a, λ_b^u and λ_b^d to the corresponding clusters' nodes. The optimal form for the sensor node update and transmission rates given in Eq. (10) and (11) can be re-written as

$$\rho_n = \sqrt{\frac{a_n}{p_n}} \quad \text{and} \quad r_n = \sqrt[3]{\frac{b_n}{\omega q_n}} \tag{12}$$

where p_n and q_n can be interpreted as *nodal update rate price* and *nodal transmission rate price* given respectively by

$$p_n = \frac{\lambda_c^a}{\mu_n^a} + \frac{\lambda_b^u}{\mu_n^u} + \frac{\lambda_b^d}{\mu^d},$$

$$q_n = \frac{1}{(\mu_n^a)^2}\frac{\lambda_c^a}{\gamma_c^a} + \frac{1}{(\mu_n^u)^2}\frac{\lambda_b^u}{\gamma_b^u} + \frac{1}{(\mu^d)^2}\frac{\lambda_b^d}{\gamma_b^d},$$

and are computed once the Lagrange multipliers are known.

Remark. We shall assume that the cluster head knows the broadcast rates μ_n^a of all $n \in \mathcal{N}_c$, while the base station knows the uplink/downlink peak rates, μ_n^u and μ^b respectively, of all $n \in \mathcal{N}$.

We summarize the algorithm as follows.

Each node computes p_n and sends $\frac{a_n}{p_n}$ to the cluster head which forwards it to the base station. The cluster head computes a cluster quantity γ_c^a that we refer to as the "spare capacity" and which is defined as follows

$$\gamma_c^a = \frac{1}{\omega}\max\left[1 - \sum_{n\in\mathcal{N}_c}\frac{1}{\mu_n^a}\sqrt{\frac{a_n}{p_n}},\ \delta\right].$$

for some small $\delta > 0$. The base station uses the quantity $\frac{a_n}{p_n}$ to compute uplink/downlink quantities, γ_b^u and γ_b^d respectively, also referred to as BS's uplink and downlink "spare capacity", and defined as

$$\gamma_b^u = \frac{1}{\omega}\max\left[1 - \sum_{n\in\mathcal{N}}\frac{1}{\mu_n^u}\sqrt{\frac{a_n}{p_n}}, \delta\right],$$

$$\gamma_b^d = \frac{1}{\omega}\max\left[1 - \sum_{n\in\mathcal{N}}\frac{1}{\mu^d}\sqrt{\frac{a_n}{p_n}}, \delta\right].$$

Note that γ_c^a depends on the update rate prices of all nodes in cluster c, while γ_b^u and γ_b^d depend on the update rate prices of all the nodes in sharing the BS. Given the nodal update prices and spare capacities, each cluster head node determines a cluster price on transmission rate given by λ_c^a/γ_c^a, while the base station determines the uplink/downlink rate transmission prices given by λ_b^u/γ_b^u and λ_b^d/γ_b^d, and then shares them with the corresponding clusters' heads. Note that the higher the spare capacity the lower the price of adopting a higher transmission rate for sensor nodes updates at node n. The transmissions' prices are then distributed from the cluster head amongst the cluster nodes. Each node in the network can now compute their own q_n. At this point, given that each node have their p_n and q_n, they update their ρ_n and r_n according to Eq.(12), then send them to the corresponding cluster heads who share them with the BS. Cluster heads update their respective prices according to

$$\lambda_c^a(t+1) = [\lambda_c^a(t) - \kappa(1 - \sum_{n\in\mathcal{N}_c}\frac{\rho_n(t)}{\mu_n^a} - \omega \parallel \mathbf{r}^c(t) \parallel_{\boldsymbol{\mu}^a,2})]^+$$

where $[x]^+ = \max[x, 0]$. Similarly, base station prices are updated at each time step (at the base station itself), as follows

$$\lambda_b^u(t+1) = [\lambda_b^u(t) - \kappa(1 - \sum_{n\in\mathcal{N}}\frac{\rho_n(t)}{\mu_n^u} - \omega \parallel \mathbf{r}(t) \parallel_{\boldsymbol{\mu}^u,2})]^+,$$

$$\lambda_b^d(t+1) = [\lambda_b^d(t) - \kappa(1 - \sum_{n\in\mathcal{N}}\frac{\rho_n(t)}{\mu^d} - \omega \parallel \mathbf{r}(t) \parallel_{\boldsymbol{\mu}^d,2})]^+.$$

The proposed algorithm is based on the natural dual decomposition approach [9] and [11], but accounts for the non-linear coupling on congested network resources. A such algorithm will naturally converge to the appropriate Lagrange multipliers, i.e., prices associated with the problem at hand.

VI. Numerical Results

We conducted various preliminary numerical evaluations to explore the algorithm's convergence and character of the resources' allocations. We considered a network with three clusters of sensing nodes sharing a single base station.

A. Convergence of the NSA Algorithm

We first show that the NSA algorithm we designed in Section V converges fairly quickly. The representative results shown in Fig.4 correspond to the case where there are 3 clusters, each having 5 nodes. The intra-cluster broadcast capacity of each node is 100 Mbps. The uplink capacity from each node to the base station is also 100 Mbps, while the downlink capacity is 100 Mbps. Under homogeneous assumptions (i.e. model parameters a_n and b_n are the same across all clusters in the network, where $a_n = 1$ and $b_n = 6$), we exhibit the convergence of the resource allocations for a single node. We note that this algorithm can in principle adapt to changing network capacities and topologies.

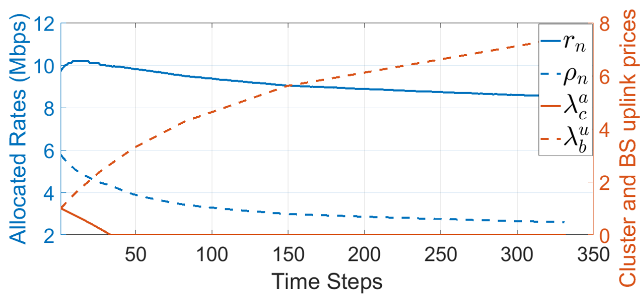

Fig. 4. Convergence of our designed NSA algorithm.

B. Increasing cluster size

Next, we studied how the nodes' update and transmission rates vary as a function of the number of sensing nodes (assuming again homogeneous conditions). For this purpose, we increase the number of nodes per cluster, N, from 2 to 10. The intra-cluster broadcast rate of each node is 100 Mbps, while the uplink and downlink capacities (from each node to BS and vice-versa) are both 50 Mbps. We plot below the optimal nodal update and transmission rates, as well as the overall network SAE. As expected, the transmission and propagation rates decrease as N increases.

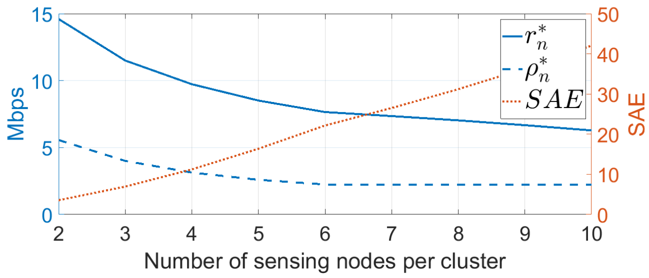

Fig. 5. Update and transmission rates vs cluster size.

C. Cluster heterogeneity in nodal SAE

Another feature we explored is the impact of giving a higher priority (or SAE sensitivity) to all the nodes in a particular cluster. This time we considered a network with two clusters with five nodes each. We assign a higher sensitivity to all the nodes of one cluster (say Cluster 1), while keeping the same weight to all the nodes of Cluster 2. An example of this scenario is when a cluster of cars is moving faster than another one. We implement this scenario as follows: Starting with $a_n = 1$ and $b_n = 6$ for all nodes in Cluster 1, we multiply them by some constant, say ζ, which we keep increasing. We then plot the rates for both clusters in function of the constant. All capacities are 100 Mbps. As expected, more rate is allocated to Cluster 1, showing that NSA optimization requires more updates and faster transmission rates to these nodes.

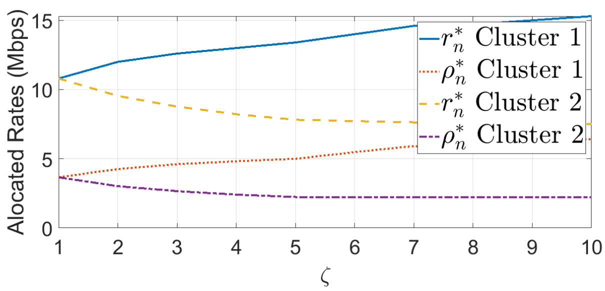

Fig. 6. Heterogeneous Setting: Leading cluster has higher SAE sensitivity.

D. Increasing the imbalance of nodes across clusters

Finally, we are interested in understanding the effect of increasing the number of nodes in one of the clusters. For this purpose, we considered a network with two clusters, 1 and 2, and a single base station. The number of nodes in Cluster 1 is kept fixed at 5, while we vary the number of nodes in Cluster 2 from 2 to 10. Below, we plot the resulting update and transmission rates, as well as the overall network SAE. We observe that our algorithm achieves fairness, which results in an equal allocation of resources among all nodes.

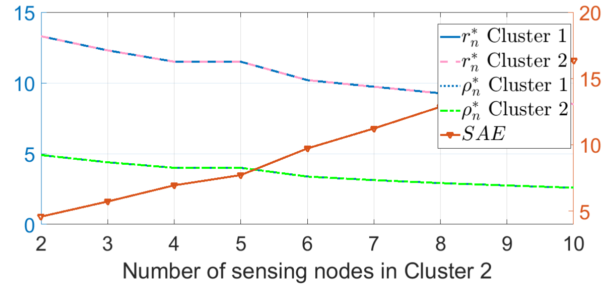

Fig. 7. Clusters with imbalanced number of nodes.

VII. Conclusion

We proposed a framework to study distributed collaborative sensing of a dynamic environment based on sharing information over limited network resources. A proposed model for situational awareness, SAE, was introduced. It is dependent on environment variability and sensor heterogeneity. The main goal was to establish key trade-offs among sensors' update and transmission rates.

We considered first a simple setting where a cluster of nodes are sharing updates over a single communication resource, which we referred to as intra-cluster broadcast, then we extended it to include multiple clusters sharing updates via a single base station, referred to as inter-cluster broadcast. We also developed a new algorithm, NSA algorithm, geared at jointly minimizing the SAE, which could optimize the allocation of resources to heterogeneous sensor nodes and time varying network capacity (and topology).

In our future work, we would like to explore more in depth a more general network setting where path routing and resource allocation decisions need to be made. This would follow from the extension of the current model to include the impact of relaying delays at the base station on SAE, along with the impact of the geographical positioning of the sensor nodes relatively to the cluster head, which will directly affect the SAE model.

VIII. Acknowledgments

We would like to thank the support of the NSF Grant ECC-1809327.

IX. Appendix

A. Proof of Lemma 1

$$
\begin{aligned}
&P\Big(\sum_{n\in\mathcal{N}} X_n \frac{r_n}{\mu_n} > 1\Big) \\
&= P\left(\sum_{n\in\mathcal{N}} X_n \frac{r_n}{\mu_n} - \mathbb{E}\Big[\sum_{n\in\mathcal{N}} X_n \frac{r_n}{\mu_n}\Big] > 1 - \mathbb{E}\Big[\sum_{n\in\mathcal{N}} X_n \frac{r_n}{\mu_n}\Big]\right) \\
&\overset{(a)}{=} P\Big(\sum_{n\in\mathcal{N}} X_n \frac{r_n}{\mu_n} - \sum_{n\in\mathcal{N}} \frac{\rho_n}{\mu_n} > 1 - \sum_{n\in\mathcal{N}} \frac{\rho_n}{\mu_n}\Big) \\
&\overset{(b)}{\le} \exp\left(-\frac{2\big(1-\sum_{n\in\mathcal{N}} \frac{\rho_n}{\mu_n}\big)^2}{\sum_{n\in\mathcal{N}} (\frac{r_n}{\mu_n})^2}\right)
\end{aligned}
$$

where ,

(a) follows from $\mathbb{E}[X_n] = \frac{\rho_n}{r_n}$, for all $n \in \mathcal{N}$.

(b) Follows from the independence of X_n's and Hoeffding upper bound.

Constraining the upper bound to be less than ϵ results in the constraint given in Eq.(8).

B. Proof of Proposition 1

One can verify that the cost function is jointly convex in $\boldsymbol{\rho}$ and $\mathbf{r}$. The congestion constraint corresponds to a convex set in both $\boldsymbol{\rho}$ and $\mathbf{r}$. We shall initially relax the constraint $\boldsymbol{\rho} \leq \boldsymbol{r} \leq \boldsymbol{\mu}$ and define the Lagrangian associated with the relaxed Problem 1.

$$\begin{aligned} L(\boldsymbol{\rho}, \boldsymbol{r}; \lambda) = -\big[\sum_{n\in\mathcal{N}} (\frac{a_n}{\rho_n} + \frac{b_n}{r_n})\big] \\ + \lambda\Bigg(1 - \sum_{n\in\mathcal{N}} \frac{\rho_n}{\mu_n} - \omega \left\|\boldsymbol{r}\right\|_{\boldsymbol{\mu},2}\Bigg) \end{aligned}$$

where $\lambda \geq 0$ is the dual variable. Let

$$f(\lambda) \triangleq \min_{\boldsymbol{\rho},\boldsymbol{r}} L(\boldsymbol{\rho}, \boldsymbol{r}; \lambda).$$

The Lagrangian dual problem is defined by

$$h \triangleq \max_{\lambda\geq 0} f(\lambda)$$

Taking the partial derivative of $L(\boldsymbol{\rho}, \boldsymbol{r}; \lambda)$ w.r.t ρ_n and r_n respectively and setting them equal to 0 gives

$$\rho_n^* = \sqrt{\frac{a_n\mu_n}{\lambda}},\ r_n^* = \mu_n \sqrt[3]{\frac{b_n \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}}{\lambda\omega\mu_n}}$$

Given optimal $\boldsymbol{r^*}$, we solve for $\left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}$

$$\begin{aligned} r_n^* &= \mu_n \sqrt[3]{\frac{b_n \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}}{\lambda\omega\mu_n}} \\ (\frac{r_n^*}{\mu_n})^2 &= \left(\frac{b_n}{\mu_n\lambda\omega}\right)^{2/3} \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}^{2/3} \\ \sum_n (\frac{r_n^*}{\mu_n})^2 &= (\sum_n (\frac{b_n}{\mu_n})^{2/3})\left(\frac{1}{\lambda\omega}\right)^{2/3} \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}^{2/3} \\ \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}^2 &= (\sum_n (\frac{b_n}{\mu_n})^{2/3})\left(\frac{1}{\lambda\omega}\right)^{2/3} \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}^{2/3} \end{aligned} \tag{13}$$

which leads to

$$\left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2} = \sqrt{\left\|\boldsymbol{b}\right\|_{\boldsymbol{\mu},\frac{2}{3}}}\left(\frac{1}{\omega\lambda}\right)^{1/2}$$

We plug-in ρ_n^* and $\left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2}$ into the constraint function and solve for λ

$$\lambda = \left[\sqrt{\left\|\boldsymbol{a}\right\|_{\boldsymbol{\mu},\frac{1}{2}}} + \sqrt{\omega \left\|\boldsymbol{b}\right\|_{\boldsymbol{\mu},\frac{2}{3}}}\right]^2$$

We finally find ρ_n^* and r_n^*,

$$\begin{aligned} \rho_n^* &= \frac{\sqrt{a_n/\mu_n}}{\sqrt{\left\|\boldsymbol{a}\right\|_{\boldsymbol{\mu},\frac{1}{2}}} + \sqrt{\omega \left\|\boldsymbol{b}\right\|_{\boldsymbol{\mu},\frac{2}{3}}}}\mu_n\ , \\ \left\|\boldsymbol{r^*}\right\|_{\boldsymbol{\mu},2} &= \frac{1}{\sqrt{\omega}}\left(\frac{\sqrt{\left\|\boldsymbol{b}\right\|_{\boldsymbol{\mu},\frac{2}{3}}}}{\sqrt{\left\|\boldsymbol{a}\right\|_{\boldsymbol{\mu},\frac{1}{2}}} + \sqrt{\omega \left\|\boldsymbol{b}\right\|_{\boldsymbol{\mu},\frac{2}{3}}}}\right), \text{ and} \\ r_n^* &= \frac{1}{\sqrt{\omega}}\left(\frac{b_n}{a_n}\right)^{\frac{1}{2}}\left(\frac{\left\|\boldsymbol{b}\right\|_{\boldsymbol{\mu},\frac{2}{3}}}{b_n/\mu_n}\right)^{\frac{1}{6}}\rho_n^* \end{aligned}$$

We re-write ρ_n^* and r_n^* in terms of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as defined in Proposition 1,

$$\rho_n^* = \frac{\alpha_n}{\parallel \boldsymbol{\alpha} \parallel_1 + \sqrt{\omega} \parallel \boldsymbol{\beta} \parallel_{\frac{4}{3}}}\mu_n, \quad r_n^* = \frac{1}{\sqrt{\omega}}\frac{\beta_n}{\alpha_n}\left(\frac{\parallel \boldsymbol{\beta} \parallel_{\frac{4}{3}}}{\beta_n}\right)^{\frac{1}{3}}\rho_n^*,$$

Next, we verify that the solution to the relaxed problem will satisfy the constraints we have relaxed. Recall that $b_n \geq a_n, \forall n \in \mathcal{N}$ and $1 \leq \omega \leq 6$ (or $e^{-72} \leq \epsilon \leq e^{-2}$). Note that

$$\frac{r_n^*}{\rho_n^*} = \frac{1}{\sqrt{\omega}}\frac{\beta_n}{\alpha_n}\left(\frac{\parallel \boldsymbol{\beta} \parallel_{\frac{4}{3}}}{\beta_n}\right)^{\frac{1}{3}} \overset{(a)}{\geq} \frac{1}{\sqrt{\omega}}\frac{\beta_n}{\alpha_n} \geq \frac{1}{\sqrt{\omega}}\sqrt{\frac{b_n}{a_n}} \geq 1,$$

where (a) follows from $\left(\frac{\|\boldsymbol{\beta}\|_{\frac{4}{3}}}{\beta_n}\right)^{\frac{1}{3}}$ being always greater than 1. The above shows that the relaxed constraint is satisfied under specific assumptions.

References

[1] S. Chen, J. Hu, Y. Shi, Y. Peng, J. Fang, R. Zhao, and L. Zhao, "Vehicle-to-everything (V2X) services supported by LTE-Based systems and 5G," *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 70–76, 2017.

[2] N. P. Antzela Kosta and V. Angelakis, *Age of Information: A New Concept, Metric, and Tool*. Foundations and Trends in Networking, 2017, vol. 12.

[3] S. Kaul, M. Gruteser, V. Rai, and J. Kenney, "Minimizing age of information in vehicular networks," in *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 8th Annual IEEE Communications Society Conference on*. IEEE, 2011, pp. 350–358.

[4] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[5] I. Kadota, E. Uysal-Biyikoglu, R. Singh, and E. Modiano, "Minimizing the age of information in broadcast wireless networks," in *54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2016, pp. 844–851.

[6] S. K. Kaul, R. D. Yates, and M. Gruteser, "Status updates through queues," in *46th Annual Conference on Information Sciences and Systems (CISS)*, March 2012, pp. 1–6.

[7] R. D. Yates and S. K. Kaul, "The age of information: Real-time status updating by multiple sources," *IEEE Transactions on Information Theory*, pp. 1–1, 2018.

[8] Y. Sun, Y. Polyanskiy, and E. Uysal-Biyikoglu, "Remote estimation of the wiener process over a channel with random delay," in *IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 321–325.

[9] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1439–1451, 2006.

[10] J. Widmer, R. Denda, and M. Mauve, "A survey on TCP-friendly congestion control," *IEEE Network*, vol. 15, no. 3, pp. 28–37, 2001.

[11] R. Srikant and L. Ying, *Communication networks: an optimization, control, and stochastic networks perspective*. Cambridge University Press, 2013.

[12] K. Nar and T. Baar, "Sampling multidimensional wiener processes," in *53rd IEEE Conference on Decision and Control*, Dec 2014, pp. 3426–3431.

[13] P. Kall, S. W. Wallace, and P. Kall, *Stochastic programming*. Springer, 1994.