Computer Networks 75 (2014) 1–17

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Operational information content sum capacity: From theory to practice ☆

Ertugrul N. Ciftcioglu [a,*], Antonios Michaloliakos [b], Aylin Yener [c], Konstantinos Psounis [b], Thomas F. La Porta [c], Ramesh Govindan [b]

[a] IBM Research, Yorktown Heights, NY 10598, United States
[b] University of Southern California, Los Angeles, CA 90049, United States
[c] Pennsylvania State University, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 23 January 2014
Received in revised form 19 June 2014
Accepted 23 August 2014
Available online 4 September 2014

Keywords:
Quality of information
Rate allocation
Network utility maximization

ABSTRACT

This paper considers Quality-of-Information (QoI) aware resource allocation policies for multiuser networks. QoI is a recently introduced composite metric which is impacted by a number of attributes of information communicated from the source(s) to the destination(s), and as such differs from traditional quality-of-service metrics considered to date. The focus of this work is defining the Operational Information Content Sum Capacity (OICC-S) of a network, achieved by the set of information attributes supported that maximize sum quality of the network. This quality is defined as a function of the information attributes provided by the source input, as well as the channel induced attributes that impact the QoI delivered to the destination(s). Optimum rate allocation to maximize the output sum quality of information and achieve OICC-S of the network for various settings is provided, and demonstrated to differ from the solution that provides maximum throughput, making QoI-awareness necessary in resource allocation. Insights arising from the analysis are provided, along with those from practical scenarios.

1. Introduction

Traditional approaches for resource allocation based on Quality of Service (QoS) perform network operations that are agnostic to the application or the information content. Such approaches may prove suboptimal for task-oriented networks where the main goal is sound decision making. Several examples for such tasks involve crowd-sourcing, participatory sensing-type applications, as well as tactical networks. To this end, a new paradigm which emphasizes the quality of information by viewing the network as an information source, and developing methods to satisfy information quality requirements at the end user is necessary.

To characterize information quality, there is growing interest in moving from traditional QoS metrics as throughput, packet delivery ratio, fairness, and delay towards new notions of quality associated with information. This effort includes introducing new attributes which characterize the value of information relevant to the specific application [1,2]. Attributes such as provenance, accuracy, precision, reliability, corroboration, credibility, age/freshness, and timeliness have been used to define the quality of information [1–4]. Event detection applications for QoI are studied in Refs. [1,5]. Recently, there have also been studies which focus on QoI-based scheduling [6–8]. In [9,7], we have optimized delivered QoI for scenarios

☆ This work was presented in part at the 14th International Conference on Information Fusion, FUSION 2011, July 2011.

* Corresponding author.
E-mail addresses: enciftci@us.ibm.com (E.N. Ciftcioglu), michalol@usc.edu (A. Michaloliakos), yener@ee.psu.edu (A. Yener), kpsounis@usc.edu (K. Psounis), tlp@cse.psu.edu (T.F. La Porta), ramesh@usc.edu (R. Govindan).

http://dx.doi.org/10.1016/j.comnet.2014.08.017
1389-1286/

with randomness in either channel conditions or traffic, focusing on a source–destination pair. In [6], we have introduced the concept of operational information content sum capacity and demonstrated initial associated theoretical results for a multisource scenario. In this work, we build on [6] to provide a comprehensive study to address QoI-aware network system optimization from both theoretical and practical aspects.

We consider the following scenario. A network is sent tasks sequentially from an end user, and users with sensing capabilities respond to these tasks. We are interested in the set of information attribute vectors that the network can support. Moreover, we identify which of these vectors of information attributes are most useful in terms of decision making associated with the task through a Quality-of-Information function. We denote the maximum sum QoI achieved by these information attribute vectors supported by the network as the *Operational Information Content Sum Capacity (OICC-S)* of the network. Proposed recently, the notion of *Operational Information Content Capacity* (OICC) is an indicator of the decision making capability that the collection of sources and links, i.e., the network can provide [2]. As such, it differs from, for instance, the Network Utility Maximization (NUM) framework where the traditional utility is a function of the flow rates [10]. While there have been recent efforts to include delay-dependent terms in the NUM framework [11,12], we take the viewpoint of optimizing of QoI metrics such as accuracy by file size adaptation at the sources. Another main difference is that while the NUM framework deals with optimal rate adaptation, we include optimization of the attributes at the source in addition to optimal rate allocation. Although the concept of QoI by itself is associated with information generated by a single source, OICC-S captures the interaction of multiple sources or flows and the physical layer they share. More specifically, we address the problem of sum quality maximization via optimal rate allocation given the application specifications and network constraints.

Among the attributes which can effect QoI and OICC-S, we focus on the effects of source-specific attributes as *accuracy* and *timeliness*.[1] Information attributes as accuracy, precision and completeness are indicators of the initial information content and the success of generating information at the sources. Timeliness, which measures the availability of information relative to the time it is needed, is related with success of network delivery. We choose accuracy and timeliness since these two attributes together capture both source and network dependent factors on quality. Accordingly, the overall OICC-S maximizing optimization framework involves both source- and link-level decisions. These sets of attributes possess a trade-off such that improving source attributes can degrade timeliness for a given network. We consider several models for QoI that depends on these two metrics.

We consider various network scenarios with the objective of maximizing the sum quality of the system, i.e., achieving the OICC-S. The main issue we address is obtaining the balance between source attributes, specifically accuracy, and timeliness for the given network, by rate allocation. QoI is a composite function of these source- and network- based attributes, hence maximization of sum quality calls for new treatment compared with the network-centric NUM framework. We first provide theoretical results for a two-user multiple access channel (MAC). For this scenario, it is well known that max weight scheduling maximizes throughput for this model by operating at one of two corner points for the MAC capacity region [13]. In contrast, here, we demonstrate that arbitrary points on the dominant face of the rate region can be optimal rate points to attain OICC-S. Next, we demonstrate that OICC-S optimizing rate allocation strategies significantly differ from throughput-maximizing rate allocation for a several canonical topologies operating with practical protocols as TDMA and CSMA/CA operating with several widely used commercial applications. We conclude, based on the analysis and the simulations that rather than focusing on maximizing the number of bits in resource allocation, QoI-aware policies are necessary to maximize the decision-making capability of a network. The organization of the paper is as follows. In Section 2, we present the basic model and QoI definitions. Next, in Section 3 we formally define the OICC-S. We provide theoretical results associated with rate allocation and information attribute optimization problems to achieve the OICC-S for different settings in Section 4. Sections 5 and 6 present scenarios with widespread applications and practical network settings. We conclude the paper in Section 7.

2. QoI: Definitions, user and application perspective

QoI is a composite, multi-dimensional metric that captures the trade-offs of several components to characterize the information ultimately delivered to the application. QoI as determined by an application is a function of both intrinsic and contextual metrics. Intrinsic metrics are those that are valued independently of the use of the information. For example, the freshness of information, i.e., its age, is a function of when the information was generated, and once delivered will have the same value regardless of the application using the information. Contextual metrics are a function of the use of the information. For instance, completeness depends on the use of information. If a photo is being used to count people in a room, it is only complete if it contains all the people in the room; if its use is to determine if at least one person is in the room, then it is complete if it shows a fraction of the room that contains one person.

Requested QoI is defined as the QoI requested by a user when issuing a task. Delivered QoI represents the QoI delivered to the user, either by retrieval of information in real-time or by retrieving information from a database.

2.1. QoI functions

QoI functions allow a requestor of information to define the relationships and trade-offs between information

[1] Other attributes such as credibility, provenance and freshness can be integrated in the framework.

metrics. For example, QoI may degrade as precision of information decreases, or improve with timeliness, i.e., as the delay in retrieving the information decreases. Consider the case of an image. By reducing the resolution of the image by resizing, timeliness is improved but accuracy and precision may be degraded. A QoI function allows the specification of the application to quantify this tradeoff. To set network controls, we need to translate QoI requirements into allocated network resources. Specifically, in order to accommodate resource assignment, we translate QoI requirements into a required network rate.

In general, the QoI derived at the end user depends on both attributes inherent to the information generated at the source, e.g., resolution, completeness, field of view, provenance, and effects of network delivery (e.g. timeliness). For instance, let us consider the application of optical character recognition (OCR), where images are sent to an end user. The accuracy requirement may be that 90 percent of characters must be decoded properly. This maps to a resolution of the page image. Obviously, the required accuracy and precision will impact how small we can make a file. Additionally, we have timeliness. The latency that is achieved is a function of the size of the file and the rate at which is used for transmission. This creates a tension (tradeoff) with accuracy and precision. Let us consider a file of s bits transmitted over a link with rate r bps. This results in an accuracy attribute of a, which is a function of s, as well as timeliness t_d as the timeliness attribute equal to $\frac{s}{r}$. We define the following QoI function as a composite function of both attributes as:

$$\mathrm{QoI}(a, t_d) = \mathrm{QoI}(a(s), t_d(s, r)). \tag{1}$$

For a given application, and given file type, there exists a specific relation between the file size of the information and its accuracy, i.e. $s(a)$ can take an arbitrary form depending on the application and the file type. We follow the natural assumption that $s(a)$ is a non-decreasing function of a for a specific type of information and application. Note that the file size s is in both a and t_d in (1), which results in a non-trivial effect on QoI.

The effect of timeliness t_d on QoI is described as a timeliness function $g(t_d)$. For this paper, we specifically consider the QoI function in the form of *accuracy*$\times$ *timeliness*, i.e.,

$$\mathrm{QoI} = a \times g(t_d). \tag{2}$$

The intuition behind this function can be explained as follows: for very large delay (bad timeliness), the information is useless for the application, hence timeliness effects override any potential accuracy benefit and results in zero QoI from the second term in (2). On the other hand, when delay is very small (timeliness is good) it has no detrimental effect on QoI and the QoI is reflected by the accuracy as the first term in (2).

For a specified application, media type and file size, we can also instantiate file size and accuracy attributes to define a function which is only dependent on the rate r, called the Quality-rate-function (QRF). In other words, we can define

$$\mathrm{QRF}_{x,y}(r) = \mathrm{QoI}(a, t_d(s, r))|_{a=a_{x,y}}, \tag{3}$$

for application x and data type y. Note that due to device capabilities or application limitations, the accuracy attributes might be further constrained. Next, we present some real-world applications where the file sizes are fixed, allowing us to demonstrate the relationship between QoI and rate.

2.1.1. Example applications and QoI functions

Here, we provide several example applications of widespread use. Specifically, we consider five applications that are of interest: image recognition, face recognition, motion detection, fingerprint recognition and optical character recognition (OCR).

First, we need to specify the QRF function for each application. The main attributes of interest are the attributes of accuracy, and timeliness.

Table 1 lists the values of accuracy attributes that result from the listed file sizes, and Max Delay, which are the deadlines for timeliness for each of the five applications that we consider. Using the parameters in Table 1 together with accuracy(size) curves from the literature and our own experiments, we obtain the curves shown in Figs. 1–5. Specifically, Fig. 1 presents the QRF function corresponding to an optical character recognition application running over a channel with a bit error rate of 10^{-3}, corrected using a Hamming code. This function has been generated on the basis of a series of image recognition tests, involving 250 separate images, in which the data transfers were conducted using the OPNET simulator. For the curve associated with the image recognition application (Fig. 5) we have used the results of experiments conducted using the open source tool Tesseract. The rest of the functions presented were created on the basis of the results presented in the following papers: for face recognition [14], for motion detection [15], and for fingerprint recognition [16].

From these applications, we can deduce two main trends for QoI:

- The QoI is zero up to some specific rate, then increases.
- The QoI tends to saturate after some rate. Hence we observe diminishing returns, where a further increase in rate does not improve QoI substantially.

The first trend means that when rate is too small, the delay is too much for the application and timeliness is very bad. The QoI remains as zero until the rate is sufficient to deliver information before the expiration deadline of the application. The second trend means that when the rate is high, timeliness is not an issue and the accuracy limits the QoI. Based on these observations, we will introduce a

Table 1
Parameter values used.

	File(image) size	Max delay	Accuracy
OCR	20 Kb	100 ms	0.67
Motion detection	–	–	0.5
Face recognition	5 Kb	20 ms	0.73
Fingerprint recognition	2 Kb	100 ms	0.07
Image recognition	20 Kb	100 ms	1

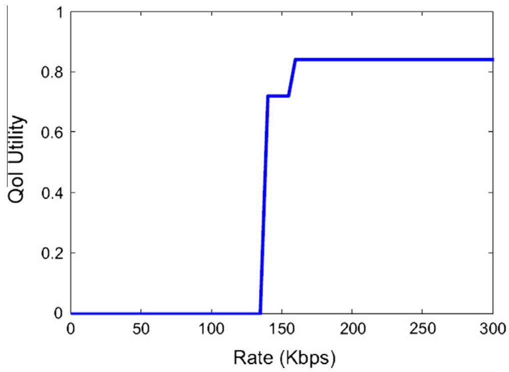


Fig. 1. OCR.

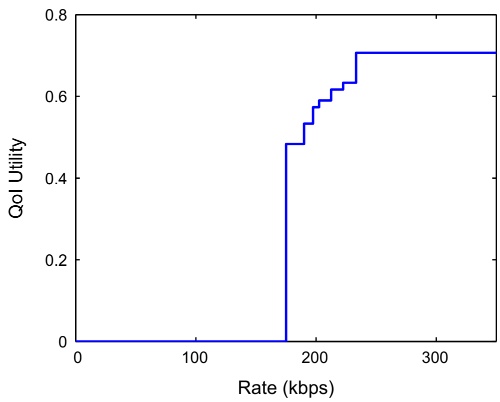


Fig. 2. Face recognition.

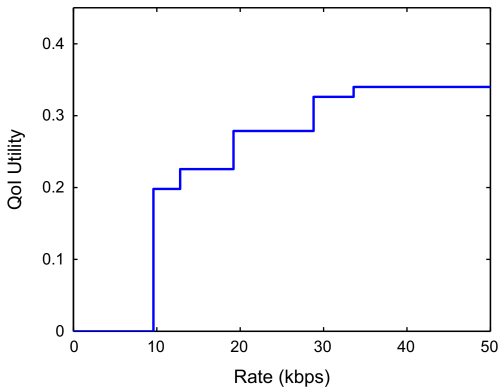


Fig. 3. Motion detection.

QoI function model with the aim of closely approximating the above characteristics, but is still amenable to mathematical analysis.

2.1.2. QoI function approximation

In this subsection, we propose a QoI function model which can be easily characterized by a small number of parameters. This function, while approximating the general trend of QoI of practical applications reasonably well, also yields for optimization methods.

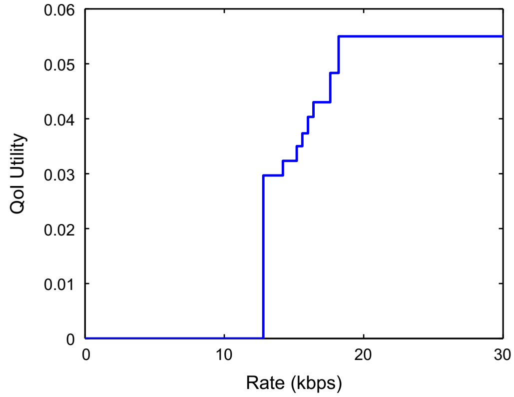


Fig. 4. Fingerprint recognition.

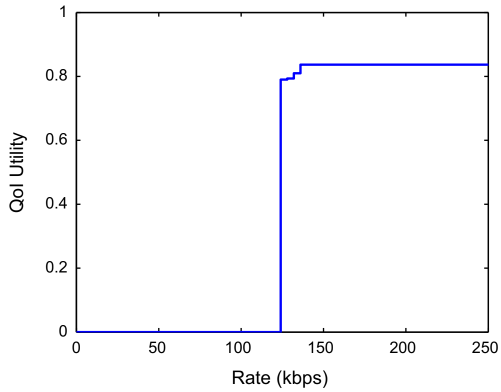


Fig. 5. Image recognition.

Next, we propose a utility function which reflects the trade-off between SAS and timeliness. More specifically, recall the QoI function from (2):

$$\mathrm{QoI}(a, t_d) = ag(t_d), \tag{4}$$

where t_d is the timeliness, i.e., delivery time of $\mathbf{q}$, and a is the accuracy attribute, and $g(t_d)$ is a timeliness function reflecting the degradation in quality due to latency. We can also express (4) in terms of a and r as follows:

$$\mathrm{QoI}(a, r) = ag\left(\frac{s(a)}{r}\right). \tag{5}$$

A function to reflect the traditional notion of timeliness could have the form that the output quality is preserved when delivered within the timeliness requirements, and reduces after some critical deadline [17]. Note that this differs from strict delay constraints which would reduce quality to zero. Piecewise linear functions can be defined for that goal. However, in order to pursue more systematic optimization methods as Lagrangian multipliers, we rather focus on smooth functions which are twice differentiable and concave within the domain of interest. As a QoI function approximating the desired property, let us consider:

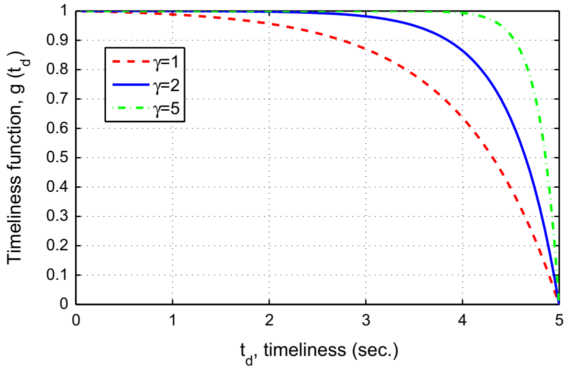


Fig. 6. Quality degradation as a function of delivery time, $D = 5$ s.

$$g(t_d) = k(\gamma, D)(1 - e^{\gamma(t_d - D)}), \tag{6}$$

for $t_d \leqslant D$. Example timeliness function curves depicting the effect of timeliness for some different parameters are illustrated in Fig. 6. Note that the general behavior of the QoI function is that it initially stays relatively unchanged for low delivery time and decays to zero as the delivery time approaches D. $D \leqslant T_{min}$ can be thought as a *maximum tolerable delay* in which the information is regarded useless afterward, and the exact behavior of the utility curve can be adjusted by varying γ. $k(\gamma, D) = \frac{1}{1-e^{-\gamma D}}$ is a normalization parameter. We also plot QRF(r) for two different file sizes, both satisfying the relationship $s(a) = \alpha a^3$ in Fig. 7. Observe that the QoI is zero until the max tolerable delay requirement is satisfied, afterward the QoI increases with rate as an artifact of improved timeliness. Both functions obey the two trends we have pointed out in the previous subsection regarding QRF functions corresponding to the example applications from practice.

3. Operational Information Content Sum Capacity (OICC-S)

We consider a scenario where tasks are issued from an end user in a tactical network. Tasks arrive with a random interarrival time greater than T_{min}. We assume that at most one task is processed by the network at any time. Information sources S_i, $i = 1, \ldots, K$ are capable of responding to this task and focus on independent events and possibly possess or generate different types of information related with the task. Moreover, each information source can respond to the task with potentially up to J applications relevant to the task, each of which will be allocated a rate. Once the resources are allocated, information available is fed into the wireless channel to the destination with a certain rate. These rates should adhere to the achievable rate region $\mathcal{C}$.

3.1. OICC-S definition

In this definition of OICC-S, there are two steps, illustrated with 2-flow cases. The descriptions will generalize to n flows or m-element information attribute vectors. The OICC-S of a given network can be derived as follows (for two information-flows $f1$ and $f2$):

For each pair (r_1, r_2) on the boundary of the achievable rate region.

- Let $(\mathbf{q}^1_{f1}, \mathbf{q}^2_{f1}, \ldots)$ (respectively for flow $f2$) be the set of all information attribute-vectors whose rates (obtained from the QoI-rate function) are less than or equal to r_1 (respectively r_2).
- Let $\mathbf{q}^j_{f1}$ (respectively $\mathbf{q}^k_{f2}$) be the information attribute vector whose QoI is the highest in the set $(\mathbf{q}^1_{f1}, \mathbf{q}^2_{f1}, \ldots)$ (respectively for $f2$).

Then, the set of all pairs $(\mathbf{q}^j_{f1}, \mathbf{q}^k_{f2})$ are candidate attribute-vector pairs for the rate pair (r_1, r_2) to attain the OICC-S of the network. Note that this set of all pairs $(\mathbf{q}^j_{f1}, \mathbf{q}^k_{f2})$ can also be equivalently identified as the set of information attribute vectors whose sum quality of information is the highest among any of the feasible $(\mathbf{q}_{f1}, \mathbf{q}_{f2})$ pairs defined above.

Ultimately, the *OICC-S* of the network is defined as the *maximum of the sum quality of information attained among all rate allocation options.*

Given a network, it is essential to optimally allocate its resources in order to achieve the OICC-S. To that end, in this subsection we first express the general formulation leading to the OICC-S of a network. First, recall that the OICC-S of a network is the maximum sum QoI attained over all rate allocation options and information attribute vectors.

Consider a network of K users, each able to select up to J applications. Furthermore assume that rates allocated to the K users are confined in the feasible set $\mathcal{C}$. Maximizing total QoI delivered to such a network is given by:

$$\max_{r_{i,j}, a_{i,j} (1 \leqslant i \leqslant K, 1 \leqslant j \leqslant J)} \sum_{i=1}^{K} \sum_{j=1}^{J} \mathrm{QoI}_j(a_{i,j}, r_{i,j}) \tag{7}$$

$$\text{s.t.} \left(\sum_{j=1}^{J} r_{1,j}, \ldots, \sum_{j=1}^{J} r_{K,j} \right) \in \mathcal{C}, \tag{8}$$

$$a_{i,j} \leqslant A_j, \tag{9}$$

where $r_{i,j}$ are the rates allocated to application j of source i, $a_{i,j}$ are the accuracy attributes of information selected from application j transmitted by source i. Recall that $\mathrm{QoI}_j(a, r)$ is the information quality attained from information from application j with rate r and accuracy a.

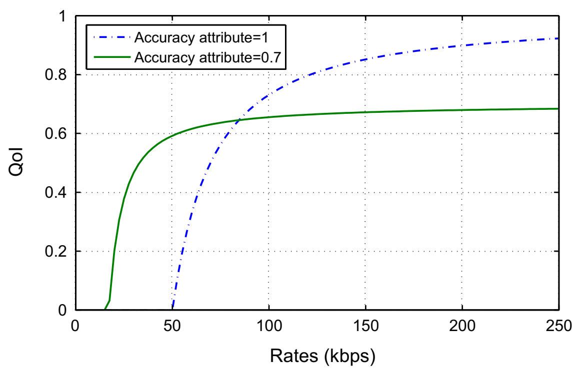


Fig. 7. QoI functions as a function of rate, different file sizes, $D = 1$ s.

A_j is the upper bound on the accuracy attribute for application j for the used information type.

Not unexpectedly, the above optimization problem is difficult to solve for large networks with arbitrary topologies. In order to solve (7), we apply iterative optimization methods. Even in cases where the joint rate allocation-optimization problem is non-convex, when the individual rate allocation and attribute optimization problems are convex, alternating maximization [18] can be applied. By doing so, the complexity can be reduced by decomposing to individual source- and network- level optimization. We next focus on simple structures that can still capture the effects of considering QoI-aware scheduling among different flows to maximize the quality of information achievable in a network.

4. Case study: The multiple access channel

In this section, we shall concentrate on uplink scenarios. More specifically, we consider the two-user Multiple Access Channel (MAC) shown in Fig. 8, and the two-user Time-Division-Multiple-Access (TDMA). The results can be readily generalized to more than two users. These constitute basic and inspiring models for analytical OICC-S characterization, which involve multiuser issues as proper rate allocation between users. This rate allocation is dependent on QoI functions and information attribute vectors from each user. We next present the set of assumptions considered in this section regarding the system model.

4.1. Transmission model

For our two-user model, transmission rates can be upper bounded by the capacity region of a Gaussian multiple access channel given by [19]:

$$r_i \leqslant W\log_2\left(1+\frac{h_iP}{N_0W}\right) = c_i, \quad i=1,2, \tag{10}$$

$$r_1 + r_2 \leqslant W\log_2\left(1+\frac{(h_1+h_2)P}{N_0W}\right) = c_s, \tag{11}$$

where r_i is the rate from S_i to destination, $\sqrt{h_i}$ denotes the channel gain from S_i to the destination node, P is the power constraint for all nodes, the $\frac{N_0}{2}$ is the noise spectral density and $2W$ is the two-sided bandwidth. We assume that channel gains are static throughout a specific task. We also assume that the time scales of interest due to timeliness requirements are large enough, along with a large operational bandwidth, allowing usage of possibly multiple codewords with sufficiently large block lengths to approach the bounds in (10) and (11) during delivery of information from the sources. Essentially, the available rate options are within a convex pentagonal region (Fig. 9), where two of the corner points correspond to different decoding orders at the destination. The significance of this rate region is that source rates are coupled via the third common constraint in (11). We emphasize that (10) and (11) constitute upper bounds for any practical protocol, as well as transmission schemes with any physical layer coding and modulation scheme.

Remark 1. While the main focus of our study in this section will entitle the rate regions given by (10) and (11), we note that the formulation we present is general enough to accommodate for alternative rate regions. For instance, let us consider *Time-Division Multiple Access (TDMA)*, where simultaneous transmission by sources are not supported. Hence we are constrained to operate in a subset of the MAC rate region. Specifically, the rate region for such a scheme is given by a triangular region (dashed line in Fig. 9), which corresponds to temporal time-sharing between two different single-user decoding points:

$$\frac{r_1}{c_2} + \frac{r_2}{c_1} \leqslant 1, \tag{12}$$

where parameters c_i are defined by (10).

4.2. Application set

We consider a class of infinitely many applications for user i, with same maximum tolerable delay D_i but different accuracy attributes $a_{i,j}$ and file sizes in order to support accuracy $a_{i,j}$ as $s_i(a_{i,j})$. For each source i, all of these applications satisfy the same QoI function (4) in terms of timeliness degradation:

$$\mathrm{QoI}_{i,j}(a_{i,j}, r) = a_{i,j}g\left(\frac{s_i(a_{i,j})}{r}\right) = a_{i,j}k_i\left(1 - e^{\gamma_i\left(\frac{s(a_{i,j})}{r} - D\right)}\right). \tag{13}$$

where γ_i is the timeliness parameter and $k_i(\gamma_i, D_i) = \frac{1}{1-e^{-\gamma_i D_i}}$ is a normalization parameter. Note that $\frac{s(a_{i,j})}{r}$ is the timeliness of the information.

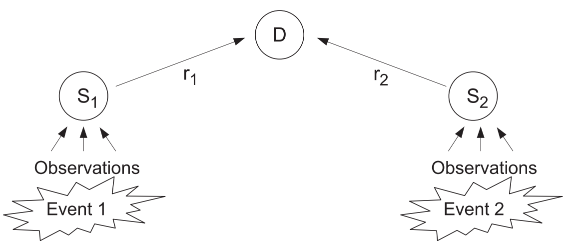


Fig. 8. Two-user MAC channel for QoI-based network.

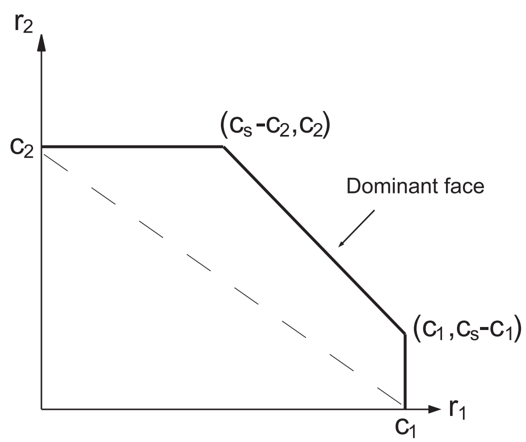


Fig. 9. Capacity region for two-user MAC channel.

For each source, we assume that at any time only one of such applications can be selected, that is we select an application with a specific file size and associated accuracy. Note that such sources still result in notably differentiable Quality-rate-functions as demonstrated in Fig. 7.

4.3. OICC-S optimization

As a result, for this model, (7) can be expressed as

$$\max_{r_1,r_2,a_1,a_2} a_1k_1\left(1-e^{\gamma_1\left(\frac{s_1(a_1)}{r_1}-D_1\right)}\right)+a_2k_2\left(1-e^{\gamma_2\left(\frac{s_2(a_2)}{r_2}-D_2\right)}\right) \quad (14)$$

$$\text{s.t. } r_i \leqslant c_i,\ i=1,2 \quad (15)$$

$$r_1+r_2 \leqslant c_s, \quad (16)$$

where r_i are the rates allocated to source i, a_i are the accuracy attributes related with QoI-vector $\mathbf{q}_{fi}$, for $i=1,2$. Timeliness parameters D_i, γ_i, and constants k_i $i=1,2$ are specific to the application.

We first note that the objective function is not jointly concave in all decision variables (r_1, r_2, a_1, a_2), so standard Karush–Kuhn–Tucker (KKT)-based optimization [20] is not readily applicable. Accordingly, we rely on iterative optimization methods based on alternating maximization [18] discussed in Section 4.3.3. Before jumping to the general solution, we next present a rate allocation problem which attains the OICC-S for special restrictions on the accuracy pairs. This problem is followed by the alternative problem of maximizing output sum QoI by information attribute adaptation for a given rate pair in Section 4.3.1. The solutions of these two problems will also constitute building blocks for the solution to achieve the OICC-S of a network for the generalized case.

4.3.1. OICC-S based rate allocation

Consider the special case the following optimization problem defined where $a_i, s_i(a_i)$, $i=1,2$, are given:

$$\max_{r_1,r_2} a_1k_1\left(1-e^{\gamma_1\left(\frac{s_1(a_1)}{r_1}-D_1\right)}\right)+a_2k_2\left(1-e^{\gamma_2\left(\frac{s_2(a_2)}{r_2}-D_2\right)}\right) \quad (17)$$

$$\text{s.t. } r_i \leqslant c_i, \quad i=1,2 \quad (18)$$

$$r_1+r_2 \leqslant c_s, \quad (19)$$

where timeliness parameters D_i, γ_i, and constants k_i $i=1,2$ are specific to the application. Hence, we are interested in the optimal rate allocation to maximize sum quality, which will in-turn define the timeliness attributes of $\mathbf{q}_{f1}$ and $\mathbf{q}_{f2}$. We first note the separability of the sum QoI in r_1 and r_2. In order to assess the applicability of standard optimization methods, we check for concavity:

$$\frac{\partial Q_r(a,r)}{\partial r} = ka\frac{\gamma s(a)}{r^2}e^{\gamma\left(\frac{s(a)}{r}-D\right)}, \quad (20)$$

Next,

$$\frac{\partial^2 Q_r(a,r)}{\partial r^2} = kas(a)\gamma\left[\frac{-2}{r^3}+\frac{-\gamma s(a)}{r^4}\right]e^{\gamma\left(\frac{s(a)}{r}-D\right)} \quad (21)$$

$$<0. \quad (22)$$

Hence the quality function is concave in rate r. We also note that the feasible region for r (MAC rate region) is a convex set.

Theorem 1. *Given accuracy attributes (a_1, a_2) of information-flows, the optimal rate allocation (r_1^*, r_2^*) is given by one of:*

1. $(r_1, r_2) = (c_1, c_s - c_1)$
2. $(r_1, r_2) = (c_s - c_2, c_2)$
3. (r_1, r_2) *on dominant face* $(r_1 + r_2 = c_s)$ *with:*

$$\frac{r_1^2}{r_2^2} = \frac{k_1\gamma_1 a_1}{k_2\gamma_2 a_2}\frac{e^{\gamma_2\left(\frac{s_2(a_2)}{r_2}-D_2\right)}}{e^{\gamma_1\left(\frac{s_1(a_1)}{r_1}-D_1\right)}}, \quad (23)$$

and the exact operating point solution can be determined by evaluating the total output QoI values. Moreover, timeliness attributes attaining the OICC-S are given by $t_{di}^* = \frac{l_i(a_i)}{r_i^*}$, *for* $i=1,2$.

Proof. In Appendix A. □

Remark 2. Throughout the analysis we made the assumption that both information-flows were served in a timely fashion, using utilities given by (6). We point out that the solution given by Theorem 1 is sufficient to cover cases where there exists information-flows which cannot be served in time. In the case where the rate region cannot support either of the flows regardless of the particular rate allocation, all candidate points will result in zero quality. As for the case where only one of the flows is supported, we note that no further improvement in sum quality can be attained by considering any additional rate pairs. This is due to the fact that the corner points of the MAC rate region given by Theorem 1 already provide full prioritization and maximum possible rate for the supported flow. □

Remark 3. Our formulation readily extends to TDMA based networks, where the only constraint on the rate region has a sum rate constraint, structurally equivalent to (11). The corner point solutions in Theorem 1 are reduced to single-user decoding solutions, while the third solution corresponds to strict time sharing between single-user decoding options. The valid solution is given by time sharing, more specifically the (r_1, r_2) satisfying:

$$\frac{r_1^2}{r_2^2} = \frac{c_2k_1\gamma_1^2 a_1}{c_1k_2\gamma_2^2 a_2}\frac{e^{\gamma_2\left(\frac{s_2(a_2)}{r_2}-D_2\right)}}{e^{\gamma_1\left(\frac{s_1(a_1)}{r_1}-D_1\right)}}. \quad (24)$$

4.3.2. OICC-S based source attribute optimization

Next, we focus on the following problem: Given fixed rate pair (r_1, r_2) on the MAC boundary, we characterize the set of information attribute-vectors $(\mathbf{q}_{f1}, \mathbf{q}_{f2})$ that attain the OICC-S.

Hence, we are interested in maximizing QoI by optimizing over accuracy attributes. Note that the incentive of possibly preferring information attribute vectors with low accuracy is that information with high accuracy may lead to excessive delay and QoI reduction due to increased file sizes and *untimely* delivery. More specifically, we consider the following problem:

$$\max_{a_1,a_2} \quad a_1k_1\left(1-e^{\gamma_1\left(\frac{s_1(a_1)}{r_1}-D_1\right)}\right)+a_2k_2\left(1-e^{\gamma_2\left(\frac{s_2(a_2)}{r_2}-D_2\right)}\right), \tag{25}$$

where rates $r_i,\ i=1,2$ are already given, and timeliness parameters $D_i,\ \gamma_i$, and constants k_i for $i=1,2$ all depend on the specific application. Note that by tracing over all $r_i,\ i=1,2$ on the MAC rate region boundary we can characterize different $(\mathbf{q}_{f1},\mathbf{q}_{f2})$ pairs.

First, we check for concavity of the quality function. Since the output QoI is separable in a_1 and a_2, we can focus on individual qualities for concavity.

Observation 1. Let $f'(a)$ and $f''(a)$ denote first- and second- order derivatives of function $f(a)$ with respect to a. The utility function is concave in a if $s(a)$ satisfies:

$$2s'(a)+as''(a)+\frac{\gamma}{r}a(s'(a))^2 \geqslant 0. \tag{26}$$

Moreover, a sufficient condition for concavity in a is $s'(a)\geqslant 0$ and $s''(a)\geqslant 0$.

Proof. In Appendix A. □

Next, we state the following theorem:

Proposition 1. *Given operating point* (r_1,r_2)*, the* $a_i^*,\ i=1,2$ *for information attribute-vectors on the OICC-S are given by the equation:*

$$a_i=\frac{r_i\left(e^{-\gamma_i\left(\frac{s_i(a_i)}{r_i}-D_i\right)}-1\right)}{\gamma_i s'(a_i)}. \tag{27}$$

Moreover timeliness attributes on the OICC-S are given by $t_{di}^*=\frac{s_i(a_i^*)}{r_i}$*, for* $i=1,2$.

Proof. The optimal point is readily obtained by equating $\frac{\partial Q_r(a,r)}{\partial a}$ to 0. □

Remark 4. For the special case where optimizing accuracy attributes from (27) exceed upper bounds A_i constrained on accuracy attributes, we can simply set the accuracy attribute solution to the upper bound A_i. This can be readily seen from the concavity of QoI in a_i and the fact that QoI is minimum (equal to 0) for $a_i=0$.

4.3.3. Joint rate allocation and QoI source attribute adaptation

In Section 4.3, we noted that the objective function in (7) is not jointly concave in the rates and accuracy attributes. On the other hand, in Section 4.3.1, we demonstrated that the objective function is concave in the rates given fixed accuracy attributes. Conditions on concavity in the accuracy attributes were also presented in Section 4.3.2. Motivated by the availability of the solutions of these two subproblems, we rely on iterative optimization. Specifically, we use alternating maximization [18] in order to solve (7) and achieve the OICC-S in the most general setting where information attribute-vectors can be adapted as well in addition to rate allocation.

The method can be described as follows:

1. Initialize $(r_1^0,r_2^0),\ (a_1^0,a_2^0)$.
2. At step $k, k>0$:
   Given (a_1^{k-1},a_2^{k-1}), maximize sum the utility by optimizing over (r_1,r_2) with solution (r_1^*,r_2^*), set $(r_1^k,r_2^k)=(r_1^*,r_2^*)$.
   Given (r_1^k,r_2^k), maximize sum the utility by optimizing over (a_1,a_2) with solution (a_1^*,a_2^*), set $(a_1^k,a_2^k)=(a_1^*,a_2^*)$.
3. Stop iteration when convergence criteria is specified.

Note that for each iteration, the rate allocation step was discussed in Section 4.3.1, and the accuracy attribute optimization was discussed in Section 4.3.2. Each iteration leads to an improved sum QoI value, approaching to the OICC-S. The final ingredient required for convergence of these iterations is boundedness of the decision variables. Note that this is already readily imposed for both the rates (r_1,r_2) by the rate region and accuracies. Hence, upper bounds could be readily included as constraints in (25) without altering the convexity of the problem.

Next, we provide a structural result regarding the OICC-S achieving resource allocation.

Proposition 2. *For linear* $s(a)$*, there exist scenarios where OICC-S is achieved by rate allocation on corner points. On the other hand, the OICC-S of the same scenarios are achieved by strict time sharing for nonlinear* $s(a)$ *relationships. Essentially, increased level of nonlinearity in the model results in the optimal rate allocation solution to deviate from linear programming-based methods.*

Proof. Consider the symmetric case where $c_1=c_2$, $\gamma_1=\gamma_2,\ D_1=D_2$ and $s_1(a)=s_2(a)$. It can be readily shown that the candidate points for solution are either the corner point or the point on the dominant face with equal time sharing, i.e. $r_1=r_2=\frac{c_s}{2}$. Let us consider the optimizing accuracy attributes and the resulting sum QoI values for the three cases. From symmetry, it is sufficient to consider only one of the corner points since they result in the same sum QoI. For given $r_i,\ i=1,2$ note that QoI-maximizing attributes satisfy the following equation:

$$g\left(\frac{s(a_i)}{r_i}\right)+a_i\left(\frac{s'(a_i)}{r_i}\right)g'\left(\frac{s(a_i)}{r_i}\right)=0. \tag{28}$$

Let us consider the case where $s(a)=a$. For this case (28) reduces to

$$g\left(\frac{a_i}{r_i}\right)+\left(\frac{a_i}{r_i}\right)g'\left(\frac{a_i}{r_i}\right)=0. \tag{29}$$

It is seen that regardless of the exact rate point, the solution is solely defined by the ratio $\frac{a_i}{r_i}$, which corresponds to the delivery time t_d. Let the solution to the equation

$g(t) + tg'(t) = 0$ be given by t_l. Accordingly, the sum-QoI maximizing accuracy attributes are given by $a_i^* = t_l r_i$. The resulting maximum sum QoI for (r_1, r_2) is given by:

$$t_l r_1 g(t_l) + t_l r_2 g(t_l) = t_l g(t_l)(r_1 + r_2), \tag{30}$$

which only depends on $r_1 + r_2$. Accordingly, any rate point on the dominant face results in equal sum QoI to both of the corner points, and the maximizing rate allocation point is not unique. Moreover, it is sufficient to consider the corner points, which would have been solutions for linear problems over the MAC rate region.

Next, let us consider a nonlinear form for $s(a)$, e.g. $s(a) = a^3$. This size-accuracy attribute relation corresponds to a case with diminishing returns in the sense that increasing the amount of bits does not linearly increase the accuracy. It also satisfies condition (26). Here, (28) reduces to

$$g\left(\frac{a_i^3}{r_i}\right) + \left(3\frac{a_i^3}{r_i}\right) g'\left(\frac{a_i^3}{r_i}\right) = 0. \tag{31}$$

Again, it is seen that regardless of the exact rate point, the solution is solely defined by the ratio $\frac{a_i^3}{r_i}$, which corresponds to the delivery time t_d. Let the solution to the equation $g(t) + 3tg'(t) = 0$ be given by t_n. Accordingly, the sum-QoI maximizing attributes are given by $a_i^* = \sqrt[3]{t_n r_i}$. The resulting maximum sum QoI for (r_1, r_2) is given by:

$$\sqrt[3]{t_n r_1} g(t_n) + \sqrt[3]{t_n r_2} g(t_n) = \sqrt[3]{t_n} g(t_n)(\sqrt[3]{r_1} + \sqrt[3]{r_2}). \tag{32}$$

Which mainly depends on $(\sqrt[3]{r_1} + \sqrt[3]{r_2})$. Now, let us consider this value for any rate pair such that $r_1 + r_2 = c_s$. Since $\sqrt[3]{r}$ is a concave function of r, from Jensen's inequality, it follows that $(\sqrt[3]{r_1} + \sqrt[3]{r_2})$ is maximized over $r_1 + r_2 = c_s$ by setting $r_1 = r_2 = \frac{c_s}{2}$. Any other rate point will result in a lower sum QoI (Fig. 10). As a result, the maximizing rate allocation point for the OICC-S is unique and is given by equal time sharing among the two corner points. We have observed that increased level of nonlinearity in the model results in the optimal rate allocation solution to deviate from linear programming-based methods. □

4.4. Numerical results

Next, we demonstrate that optimal rate allocation can be different from a corner point of the rate region for various scenarios.

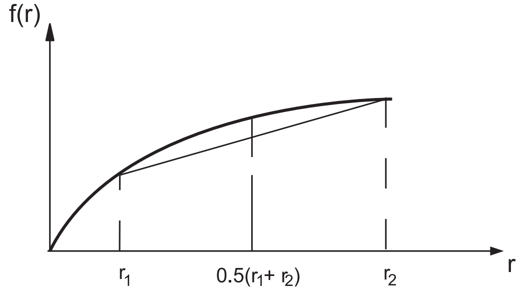


Fig. 10. Concave function of rates.

First, consider the scenario with information types, information attribute-vectors, timeliness properties, link qualities and device capabilities characterized by timeliness parameters $\gamma_1 = 3$, $\gamma_2 = 1$, maximum tolerable deadlines of $D_1 = 600$ ms, $D_2 = 750$ ms, and rate bounds $c_1 = 212$ Kbps, $c_2 = 142$ Kbps, $c_s = 259$ Kbps. We assume that $s(a) = a^3 \times 10^5$ is the relationship between file size and accuracy attribute. This corresponds to a case where accuracy achieved is a concave function of file size. The intuition is that utility gains are diminishing in return; after some level the accuracy and the effect to QoI tends to saturate. Moreover, it satisfies the condition to preserve concavity of QoI in a_i given by (26). We present the OICC-S offered by the network as a function of the accuracy attributes in Fig. 11. Fig. 12 demonstrates the optimizing r_1, i.e. rate from source 1 to achieve the corresponding OICC values in Fig. 11. In essence, these two figures demonstrate that optimal rate allocation and the resulting OICC-S greatly depends on the information attributes, and for many cases time-sharing is the optimum rate allocation choice. For this scenario, the OICC-S is 1.049 achieved by $(r_1, r_2) = (133\text{ Kbps}, 126\text{ Kbps})$ and $(a_1, a_2) = (0.67, 0.66)$. Note that the optimizing rate point is achieved by time-sharing.

Next, we consider the OICC-S offered by the network as a function of the accuracy attributes with identical parameters except $s_i(a_i) = a_i \times 10^5$. We do not explicitly depict the OICC-S and optimum rate allocation as a function of accuracy attributes since they are in general similar to Figs. 11 and 12. For this scenario the OICC-S of the network is 0.576, achieved by $(r_1, r_2) = (117\text{ Kbps}, 142\text{ Kbps})$ and $(a_1, a_2) = (0.42, 0.58)$. Note that the optimizing rate point in this scenario is the corner point where information from source 2 is decoded later. While this outcome is considerably different from the first scenario, we point out that the intuition is in line with Proposition 2. Specifically, the higher level of nonlinearity in the objective function due to the $s(a)$ relationship in the first scenario tends to cause the solution to deviate more from linear programming based solutions, i.e., corner points of the MAC rate region.

Hence, in many scenarios a simplified policy only focusing on corner points which was sufficient for traditional QoS-based objectives cannot provide the network with the maximum QoI, i.e., attained OICC-S for the available information at hand or equivalently decision making capability.

Finally, we also present analytical results for a TDMA-based two-user network. Consider the two-user network with single user maximum rates $c_1 = 300$ Kbps and $c_2 = 150$ Kbps. The overall rate region is the triangular region defined by the closure of $(0, 0)$, $(300, 0)$ and $(0, 150)$. In this scenario, contrary with the MAC rate region, the total sum rate of the time sharing options is not constant, and maximum sum throughput from the network is attained by always scheduling the user with higher rate, in this case a rate allocation of $(300, 0)$.

We consider the same set of QoI functions and attributes as the first scenario MAC example with diminishing returns, i.e. $s(a) = a^3 \times 10^5$. Overall, we observe that while sum throughput maximizing rate allocation results in a QoI of 0.66, the OICC-S of the network is 0.996 is attained by the rate points $(178.5, 60.75)$. In other words, by QoI-aware

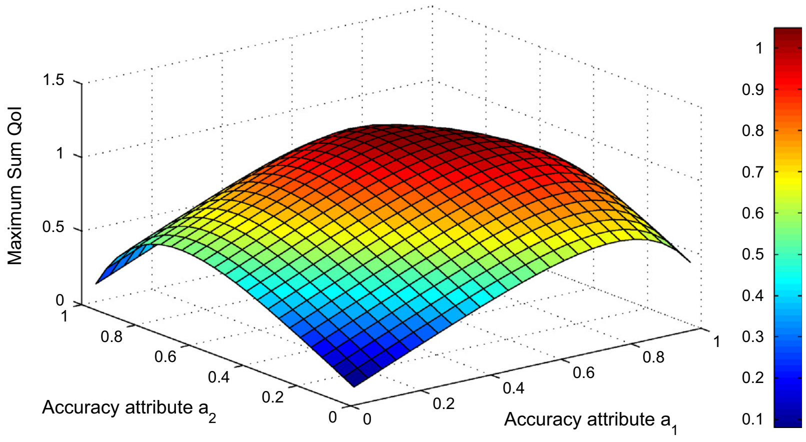


Fig. 11. Maximum Sum QoI for varying a_1 and a_2, $s(a) = a^3 \times 10^5$.

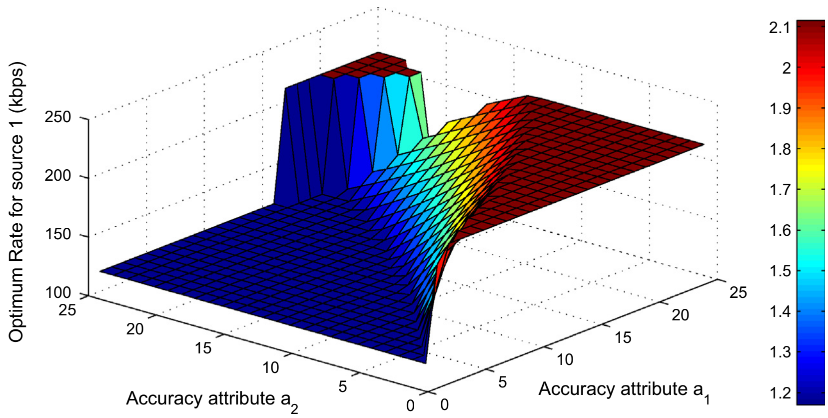


Fig. 12. Optimal r_1 for varying a_1 and a_2, $s(a) = a^3 \times 10^5$.

resource allocation, OICC-S is increased by 50 percent with a total throughput of only 239.25 compared with 300. Again, the optimizing rate points and resulting Sum QoI depends on the accuracy pairs. While we do not depict the Sum QoI as the trends are similar to Fig. 11, as presented in Fig. 13, for vast majority of the attribute pairs, sum QoI is maximized by strict time-sharing, implying throughput-maximizing algorithms do not maximize sum QoI.

5. Case study: TDMA based network with multiple applications

In the previous section, we have analytically demonstrated the necessity for QoI-aware scheduling and optimization, as the optimal solutions can deviate from traditional QoS-based network solutions. While this is a fundamental result, recall from Section 4 that it is not tractable to theoretically analyze any given network scenario and application in detail.

To that end, in this section to make the analysis tractable, we relax the constraints on the application characteristics. Specifically, in this section rather than the analytical model built for QoI functions, we consider the example applications from Section 2.1.1. While the MAC considered in Section 4 provides fundamental upper bounds on the rate regions, many commercial applications operate on more simpler protocols as TDMA, which is able to provide a subset of the MAC rate region. Accordingly, we pursue our study with the real-world applications with the TDMA protocol in this section. Suppose we have two mobile nodes and a stationary base-station to which the mobile nodes send the data of their assigned tasks.

First, we define QoI-to-Rate Functions (QRF) for cases in which multiple tasks are serviced by multiple nodes in a network. This may represent a single user requesting multiple modes of information, possibly at the same time if the rate can support it.

Suppose T tasks/applications are assigned to a node, where each task $i, 1 \leqslant i \leqslant T$, has a QRF function $Q_i(r)$. The multi-application QRF function, $Q(r)$, is defined as follows:

$$Q(r) = \max_{\mathbf{r}} \quad \sum_{i=1}^{T} Q_i(r_i)$$
$$\text{subject to}: \quad \sum_{i=1}^{T} r_i \leqslant r, \tag{33}$$

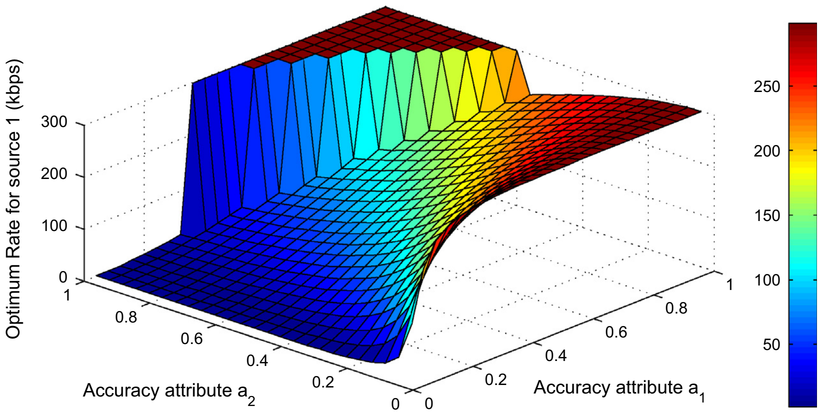


Fig. 13. Optimum r_1 for varying a_1 and a_2, TDMA.

Specifically, let us define the two sets of multi-application tasks: (T1) {motion detection, OCR, face recognition} and (T2) {fingerprint recognition, image recognition, face recognition}, where the individual application QRFs where introduced in Section 2.2.1. T1 and T2 have the multi-application QRF function as given in Fig. 14a and b.

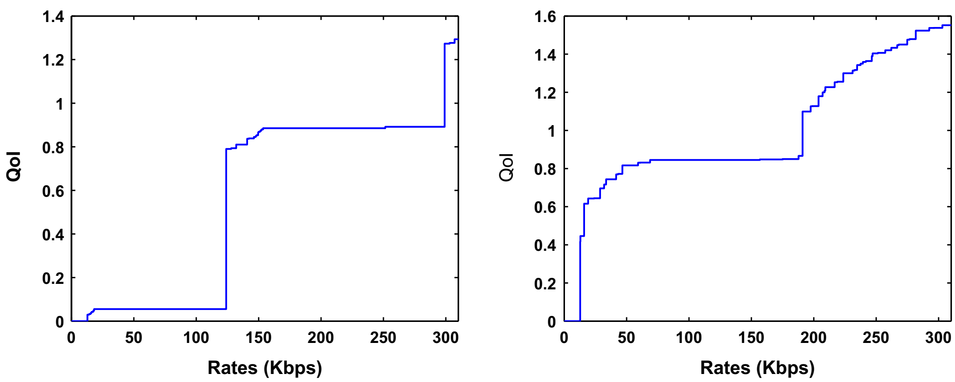
The task sets (T1) and (T2) defined above are assigned to the first and second node, respectively. Using a centralized scheduler in the base-station (e.g. TDMA), this network can achieve the triangular rate region suggested by Fig. 9.

The idea here is that nodes may concurrently use a number of applications, and different nodes may be using different applications as well as different number of applications. With the possibility of simultaneously allocating rates of each source among the multi-application tasks, the corresponding *multi-application QRF* functions associated with the two task sets are shown in Fig. 14a and b, respectively. In general, if the rates at which applications begin to be useful are distant enough from each other in the corresponding QRF functions, we expect to have in the resulting multi-application QRF function as many "regions" as the number of applications.

We first consider a network where the triangular rate region is defined by single user capacities $c_1 = 300$ kbps and $c_2 = 150$ kbps. It can be readily seen that a traditional QoS-based scheduler which aims to maximize sum throughput in the system would allocate the rate point $(300, 0)$. On the other hand, in this section the aim is to maximize the OICC-S attained from the network for the two task sets T_1 and T_2. It can be readily seen that scheduling only one user does not provide the OICC-S maximizing solution. Particularly, we observe that the OICC-S of the network is 1.7301, attained by the rate pair (153.6 Kbps,68.8 Kbps), which has a total throughput of 222.4 Kbps, significantly lower compared with 300 Kbps resulting from the max throughput scheduler which gives a sum QoI of 1.2733.

Hence, we have demonstrated that QoI-aware scheduling can result in very significant gains compared with QoS-based scheduling for a scenario with commonly used protocols and applications.

Next, in order to provide a more comprehensive demonstration of the benefits of OICC-aware scheduling, let us fix the maximum capacity allocated to one of the flows, and vary the capacity that can be allocated to the others. This results in different scenarios described by several rate



(a) Task 1 (T1). (b) Task 2 (T2).

Fig. 14. Multiapplication QRF functions.

regions. For instance, we consider the scenario where the capacity from source 2 is fixed at 150 kbps, while the capacity for link the link from source 1 is varied. From Figs. 15 and 16, we observe that the OICC-S aware scheduler consistently outperforms traditional schedulers significantly in terms of sum QoI, with less bits transmitted as demonstrated.

6. Case studies: Canonical and arbitrary network topologies with CSMA/CA

Having presented QoI characteristics of several real-world applications and how multiple such applications can be scheduled among two sources, it is interesting to study the behavior of the OICC-S achieving resource allocation in more complicated scenarios, which have been previously used in order to understand the complications posed by wireless multi-hop networks in realizing scheduling and congestion control schemes. We have taken a number of such canonical scenarios as well as a random/arbitrary topology scenario and computed their respective OICC-S values and resource allocation solutions. The computation of the rate region of these scenarios has been done assuming the classical 802.11 CSMA/CA airtime contention algorithm and the model developed in [21]. The model takes into account all protocol components and topological effects and asymmetries, along with data payload (DATA) and Request-to-Send (RTS)/Clear-to-Send (CTS) packet collisions, exponential back-off times, virtual and physical carrier sensing, channel losses due to fading, and inter-dependence between both neighboring and non-neighboring edges. Based on the analytical computation of the collision probabilities that is developed for two node topologies and extended in effectively any multi hop topology, the authors are able to compute the service time for every edge and thus the achievable rate region of the topology. The model has been verified through simulation and the analytical results are used here for characteristic topologies presented in [22].

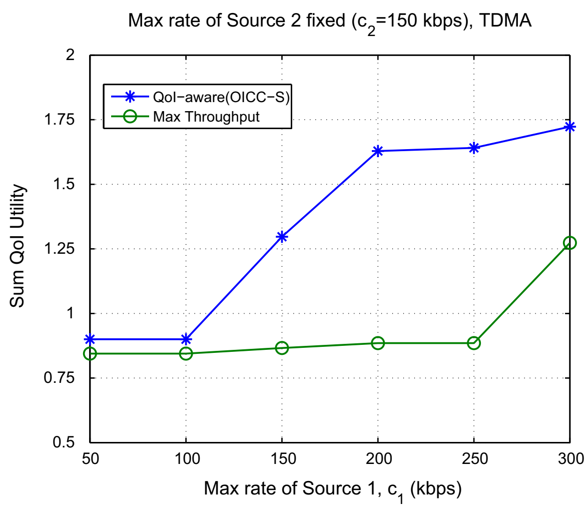


Fig. 15. Sum QoI for different schedulers.

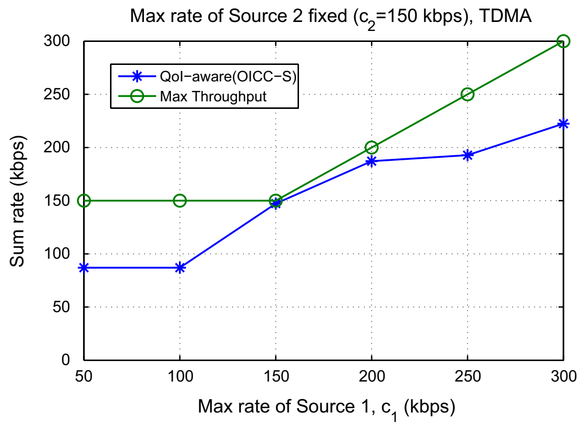


Fig. 16. Sum rates for different schedulers.

In order to provide diversity in scenarios analyzed, we consider both a large number of different topologies, and for each specific topology we consider two different cases with different link rates (1 Mbps and 11 Mbps). We refer readers to [22] for rate regions of these topologies for these different link rates. Unless otherwise specified, our discussions regarding comparison of different schedulers are carried out assuming data rate of 1 Mbps.

Chain: The first such scenario is a chain topology, for which the graphical representation is presented in Fig. 17. This presents two long flows which share a possibly infinite chain network. The boundary of the rate region of this scenario is simply a line and the rate region is symmetric. The data of the upper flow corresponds to the first task set, T1, defined previously and the data of the lower flow correspond to the second task set, T2. The resulting OICC-S maximizing rate allocation selects (124 Kbps, 19.2 Kbps) for the two flows and attains an OICC-S of 1.4332 as compared to the 0.8451 that can be achieved using a max rate allocation.

Chain with Two Interfering Short Flows (C2SF): Presented in Fig. 18, this topology is similar to chain, but here one long flow interferes with multiple short flows. By symmetry, the two short flows will achieve approximately the same rate for any scheme. In this scenario, the short flows carry packets for tasks chosen from the first task set while the long flow carries packets for tasks chosen from the second task set. OICC-S aware scheduling results in 14 percent improvement in Sum QoI.

Chain-cross: The next scenario considered is a chain-cross topology, in Fig. 19. In this case, allotting a certain rate to the long flow decreases the rate of the short flows significantly. This topology is similar to CS2F except that short flows around node 2 do not interfere with each other. Note that in our rate assignment, all short flows operate at the same rate and, for the same application, experience the same QoI. We observe an improvement of 26 percent in sum QoI compared with traditional schedulers.

Flow in the Middle (FIM): As presented in Fig. 20, in this topology one congested middle flow interferes with two outer flows. Outer flows do not interfere with each other. Each flow experiences different level of interference, with the middle flow experiencing more interference from

Fig. 17. Chain topology.

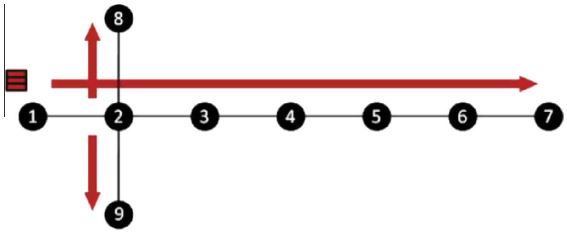


Fig. 18. Chain topology with two short flows.

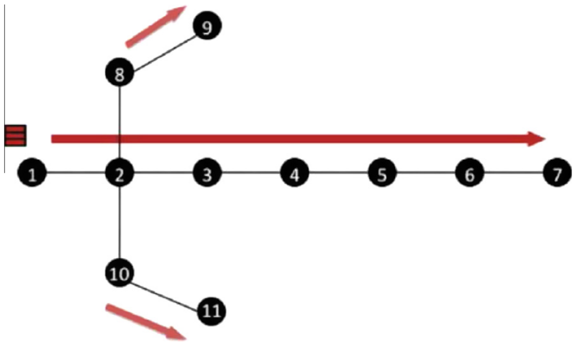


Fig. 19. Chain-cross topology.

the outer flows. OICC-aware resource allocation results in 26.5 percent improvement in sum QoI with less bits transmitted, with 387.5 kbps assigned to outer flows and 156.8 kbps assigned to the middle flow.

Stack: Fig. 21 presents the stack scenario, where the outer and inner flows use, two different QRF functions, corresponding to the first and second task sets respectively. The main difference compared with the FIM topology is that each flow goes to two hops instead of one. Despite the symmetric shape of the rate region, the OICC-S maximizing rate allocation is made asymmetric due to the different QRF functions used by the flows. Specifically, an OICC-S of 2.5851 is attained with a rate allocation of 153.6 Kbps for the outer flows and 133.1 Kbps for the inner flow, for comparison the max rate allocation achieves an OICC-S of 1.5518.

Fork: The fork topology is similar to the flow in the middle and stack topologies, but here the middle flow interferes with three non-interfering flows instead of just two (Fig. 22). OICC-aware scheduling results in 8 percent and 32 percent improvements compared with QoS-aware schedulers for different link rates.

We observe from Figs. 23 and 24 that for all combinations of topologies and link rates, the OICC-S aware resource allocation notably improves sum QoI of all flows. Furthermore, as demonstrated in Figs. 25 and 26 this improvement in sum QoI is attained with significantly less

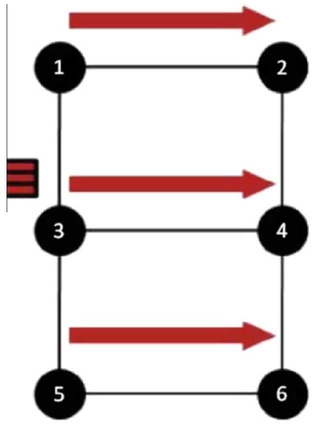


Fig. 20. Flow-in-the-middle topology.

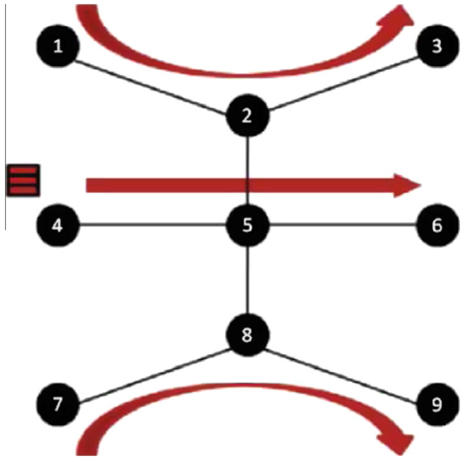


Fig. 21. Stack topology.

bits transmitted. These results confirm that QoI-aware scheduling results in much more efficient resource allocation tailored to the specific application.

Arbitrary Topology: Finally, using the ns-2 [23] network simulator, we computed the rate region of a random multi hop 802.11 CSMA topology with two arbitrary flows and run the OICC-S aware scheduler for that topology. The random topology was created scattering 10 nodes in a 500 m × 500 m field using the ns-2 scene generator tool (*setdest*) with the default seed. The default characteristics of the 914 MHz Lucent WaveLAN DSSS radio interface were used as can be found in the simulator. A basic rate of 1 Mbps along with a data rate of 1 Mbps or 11 Mbps was chosen for the MAC layer. We have observed that for a data rate of 1 Mbps, OICC-aware resource allocation results in a Sum QoI of 1.6697 with sum rate of 195 Kbps, while the

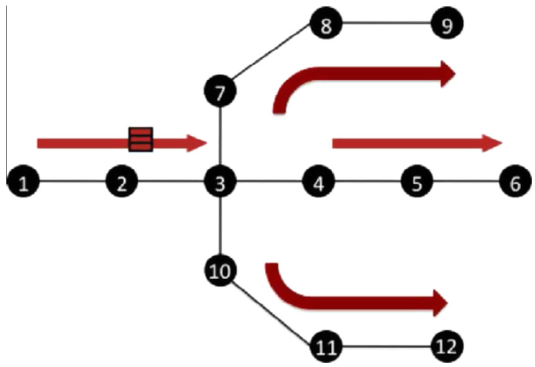


Fig. 22. Fork topology.

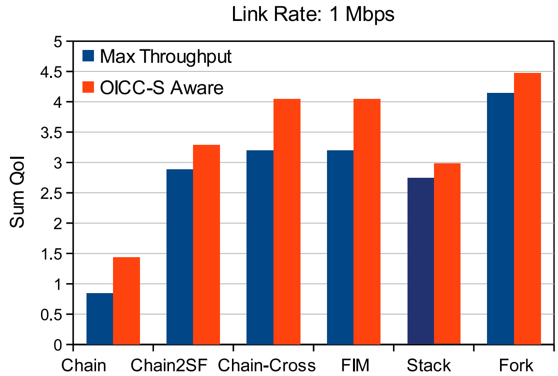


Fig. 23. Sum QoI for different topologies.

throughput maximizing only achieves a sum QoI of 0.885 despite the 237.4 Kbps. Hence we observe almost 90 percent improvement with 17 percent less bits. For 11 Mbps, the OICC-S is 2.275 attained by a sum rate of 398.3 Kbps, while the throughput maximizing scheduler results in a sum QoI of 1.5983 with 427 Kbps, again confirming the significant gains of QoI-aware schedulers despite requiring less number of bits.

6.1. OICC region of the network

Finally, for a sample topology, specifically flow-in-the-middle, we demonstrate the region of achievable individual QoI values of different flows. We call this region the *OICC region* of the network. This quantity is reminiscent of the well known concept of capacity region, or achievable rate region, which is used in traditional network theory to express the vector of rates that can be supported by the network, see, for example [21,24]. The fundamental difference here is that we are not merely interested in the rates that can be supported. Instead, we care about the information content that can be supported, i.e. in the quality of the information that can be transferred and the resulting quality of experience for the user.

Fig. 28 highlights some properties of the OICC region of the network with rate region given in Fig. 27 under study which we discuss below. Note that the concept of an OICC region is quite complex and general, and it is beyond the scope of this paper to provide a full exposition of the concept.

- The region is composed of $(n_1 + 1) \times (n_2 + 1)$ discrete regions, where n_1 and n_2 are the number of distinguishable applications available to the first and the second user respectively. For example, in Fig. 28, region 1 corresponds to the image recognition task of the first user and the face recognition task of the second user, region 2 corresponds to the case in which the second user does not achieve any QoI because the user gets no rate or too small rate, whereas the first user does achieve some QoI.

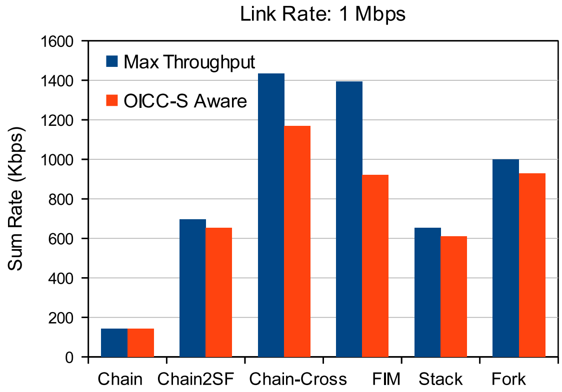


Fig. 25. Sum throughput for different topologies.

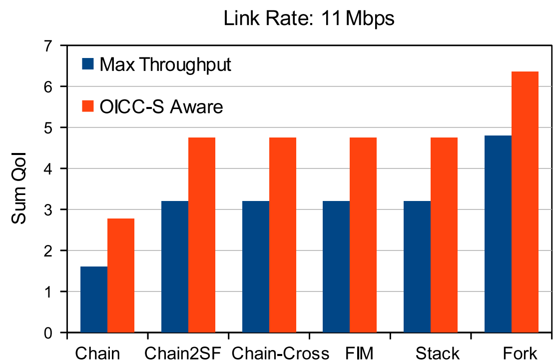


Fig. 24. Sum QoI for different topologies.

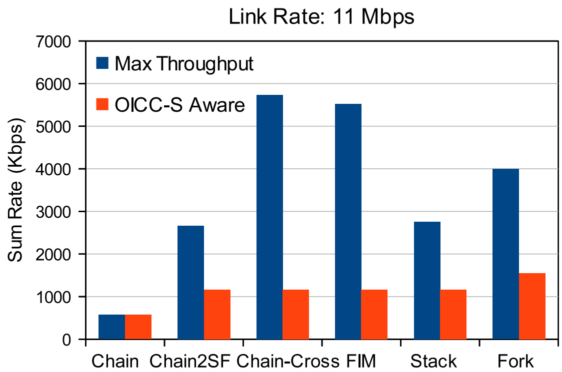


Fig. 26. Sum throughput for different topologies.

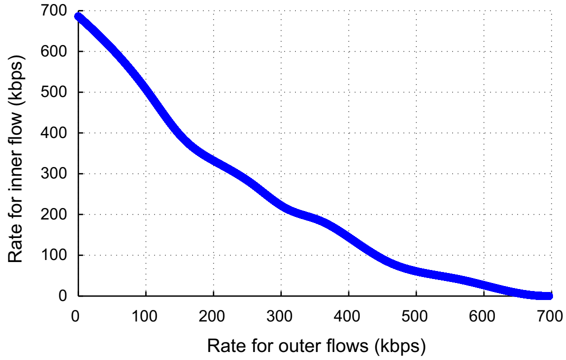


Fig. 27. Rate region for flow-in-the-middle topology.

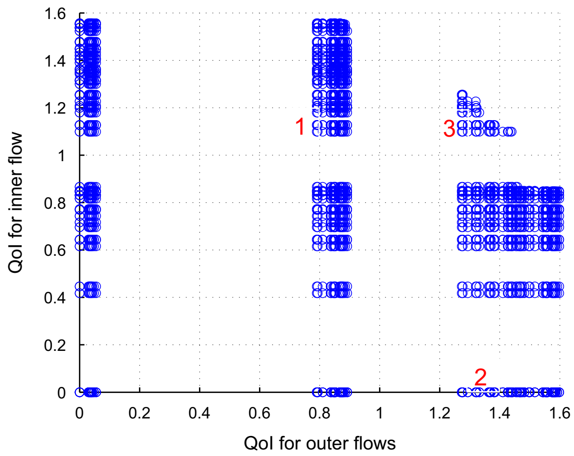


Fig. 28. The QoI region for FIM topology.

- Region 3 in the figure clearly shows the effect of the rate region on the shape of the OICC region. The upper right corner is constrained because the corresponding rate vectors are not achievable.

Remark 5. While recently protocols which combine CSMA and TDMA as hybrid MAC protocols have also been developed ([25,26] and references therein), the main specifics of extending our study to such protocols would be the new rate regions. On the other hand, the case studies presented in this paper demonstrate benefits of QoI-aware scheduling for a diverse set of different rate regions, as information theoretic, TDMA-based and CSMA-based with numerous topologies.[2] These rate regions encompass both convex and non-convex regions with different shapes and magnitudes, implying that the benefits of QoI-aware networking would readily follow for different protocols as hybrid MAC as well.

[2] While we have not explicitly presented the rate regions of the canonical topologies to keep the paper focused, we refer readers to [21,22] for rate regions of these topologies.

7. Conclusion

In this paper, we propose methods for QoI based evaluation in multiuser networks. We characterize the maximum sum output QoI provided by information-vectors supportable by the network as the OICC-S. For OICC-S formulation, we focus on the effect of network delivery and timeliness on information with specific accuracy attributes. We first theoretically characterize rate allocation schemes in order to attain OICC-S for the most basic multiuser network model, specifically a two-user MAC. Next, we provide OICC-S optimizing rate allocation solutions for several realistic QoI functions with the commonly applied TDMA and CSMA/CA protocols to further demonstrate the necessity and merits of QoI-aware networking. Results from both theoretical and practical viewpoints reveal that QoI-aware networking calls for optimization and resource allocation beyond traditional QoS-based methods.

Acknowledgements

Research was sponsored by the U.S. Army Research Laboratory under the Network Science Collaborative Technology Alliance, Agreement Number W911NF-09-2-0053. The views and conclusions contained in this document are those of the author(s) and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

Appendix A

Proof of Theorem 1. Let us introduce Lagrange multipliers λ_1, λ_2, λ_3, all greater than or equal to 0, for constraints (18) and (19). The Lagrangian can be expressed as:

$$L(r_1, r_2, \lambda_1, \lambda_2, \lambda_3) = -\sum_{i=1}^{2} a_i k_i \left(1 - e^{\gamma_i\left(\frac{s_i(a_i)}{r_i} - D_i\right)}\right) + \sum_{i=1}^{2} \lambda_i (r_i - c_i) + \lambda_3 (r_1 + r_2 - c_s). \quad (34)$$

KKT conditions dictate:

$$-k_i a_i \frac{\gamma_i s_i(a_i)}{r_i^2} e^{\gamma_i\left(\frac{s_i(a_i)}{r_i} - D_i\right)} + \lambda_i + \lambda_3 = 0, \quad (35)$$

$$\lambda_i (r_i - c_i) = 0, \quad i = 1, 2 \quad (36)$$

$$\lambda_3 (r_1 + r_2 - c_s) = 0. \quad (37)$$

Hence we have:

$$k_i a_i \frac{\gamma_i s_i(a_i)}{r_i^2} e^{\gamma_i\left(\frac{s_i(a_i)}{r_i} - D_i\right)} = \lambda_i + \lambda_3, \quad i = 1, 2. \quad (38)$$

Note that these imply that $\lambda_1 + \lambda_3 > 0$ and $\lambda_2 + \lambda_3 > 0$. First, assume $\lambda_3 = 0$. Then, $r_1 + r_2 < c_s$ and $\lambda_1 > 0, \lambda_2 > 0$ should be satisfied leading to $r_1 = c_1$, $r_2 = c_2$ but this combination is not feasible ($c_s < c_1 + c_2$). Hence it is required that $\lambda_3 > 0$, and accordingly we have $r_1 + r_2 = c_s$.

As for λ_1 and λ_2, we have the option that only one of them is positive, which would correspond to one of the corner points of the rate region. The other option is that when $\lambda_1 = \lambda_2 = 0$, which implies that $r_1 < c_1$ and $r_2 < c_2$. Along with $r_1 + r_2 = c_s$, this results in an operating point on the dominant face of the rate region (which is achieved by strict time-sharing between the two corner points corresponding to different decoding order at the receiver). From (38), with $\lambda_1 = \lambda_2 = 0$ we have

$$k_1\gamma_1 a_1 \frac{s_1(a_1)}{r_1^2} e^{\gamma_1\left(\frac{s_1(a_1)}{r_1}-D_1\right)} = k_2\gamma_2 a_2 \frac{s_2(a_2)}{r_2^2} e^{\gamma_2\left(\frac{s_2(a_2)}{r_2}-D_2\right)}, \tag{39}$$

leading to Eq. (23). In other words, the operating point is the point on the dominant face satisfying (23). The specific point will depend on multiple parameters, including accuracy attributes and timeliness parameters. □

Proof of Observation 1. Since

$$\frac{\partial^2 Q_r(a,r)}{\partial a^2} = -k\left[\frac{\gamma}{r}(2s'(a)+as''(a)+\frac{\gamma}{r}a(s'(a))^2)\right]e^{\gamma\left(\frac{s(a)}{r}-D\right)} < 0, \tag{40}$$

the quality function is also concave in accuracy a if (26) is satisfied. The sufficient condition stated is readily shown to satisfy this requirement. □

References

[1] S. Zahedi, C. Bisdikian, A framework for QoI-inspired analysis for sensor network deployment planning, in: International Workshop on Performance Control in Wireless Sensor Networks, PWSN, Austin, TX, 2007.

[2] A. Bar-Noy, G. Cirincione, R. Govindan, S. Krishnamurthy, T.F. LaPorta, P. Mohapatra, M. Neely, A. Yener, Quality-of-information aware networking for tactical military networks, in: Third International Workshop on Information Quality and Quality of Service for Pervasive Computing, in Conjunction with IEEE Percom 2011, Seattle, WA, 2011.

[3] B. Liu, P. Terlecky, A. Bar-Noy, R. Govindan, M.J. Neely, Optimizing information credibility in social swarming applications, in: IEEE INFOCOM 2011 Mini-Conference, Shanghai, China, 2011.

[4] J. Edwards, A. Bahjat, Y. Jiang, T. Cook, T.F. LaPorta, Quality of information-aware mobile applications, Pervasive Mobile Comput. 11 (2014) 216–228.

[5] Z.M. Charbiwala, S.Z.Y. Kim, Y. Cho, M.B. Srivastava, Toward quality of information aware rate control for sensor networks, in: Fourth International Workshop on Feedback Control Implementation and Design in Computing Systems and Networks, San Francisco, CA, 2009.

[6] E.N. Ciftcioglu, A. Yener, R. Govindan, K. Psounis, Operational information content sum capacity: formulation and examples, in: Proc. 14th Conference on Information Fusion, Chicago, IL, 2011.

[7] R. Urgaonkar, E.N. Ciftcioglu, A. Yener, M.J. Neely, Quality of information aware scheduling in task processing networks, in: Proc. The 7th International Workshop on Resource Allocation and Cooperation in Wireless Networks (RAWNET), in Conjunction with IEEE WiOpt 2011, Princeton, NJ, 2011.

[8] E.N. Ciftcioglu, A. Yener, M.J. Neely, Maximizing quality of information from multiple sensor devices: the exploration-exploitation tradeoff, IEEE J. Sel. Top. Signal Process. 7 (5) (2013) 883–894.

[9] E.N. Ciftcioglu, A. Yener, Quality-of-information aware transmission policies with time-varying links, in: Proc. MILCOM 2011, Baltimore, MD, 2011.

[10] F. Kelly, Charging and rate control for elastic traffic, Eur. Trans. Telecommun. 8 (1997) 33–37.

[11] H. Xiong, R. Li, A. Eryilmaz, E. Ekici, Delay-aware cross-layer design for network utility maximization in multi-hop networks, IEEE J. Sel. Areas Commun. 29 (5) (2011) 951–959.

[12] A. Szwabe, P.M.P. Walkowiak, Delay-aware NUM system for wireless multi-hop networks, in: Wireless Conference 2011 – Sustainable Wireless Technologies (European Wireless), 11th European, Vienna, Austria, 2011.

[13] R.A. Berry, E.M. Yeh, Cross-layer wireless resource allocation – fundamental performance limits for wireless fading channels, IEEE Signal Process. Mag. 21 (5) (2004) 59–68.

[14] Y. Yan, R. Muraleedharan, X. Ye, L. Osadciw, Contourlet based image compression for wireless communication in face recognition system, in: IEEE International Conference on Communications, 2008. ICC '08, 2008, pp. 505–509.

[15] J.-F. Yang, S.-C. Chang, C.-Y. Chen, Computation reduction for motion search in low rate video coders, IEEE Trans. Circ. Syst. Video Technol. 12 (10) (2002) 948–951.

[16] A. Mascher-Kampfer, H. Stogner, A. Uhl, Comparison of compression algorithms' impact on fingerprint and face recognition accuracy, in: Visual Computing and Image Processing, VCIP, 2007.

[17] A.M. Wyglinski, M. Nekovee, Y.T. Hou, Cognitive radio communications and networks: principles and practice, Academic Press, 2009.

[18] I. Csiszar, G. Tusnady, Information geometry and alternating minimization procedures, Stat. Decis. Suppl. Issue (1) (1984) 205–237.

[19] T.M. Cover, J.A. Thomas, Elements of information theory, John Wiley and Sons, New York, NY, 1991.

[20] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, Nonlinear programming, third ed., John Wiley and Sons, Hoboken, NJ, 2006.

[21] A. Jindal, K. Psounis, The achievable rate region of 802.11-scheduled multi-hop networks, IEEE/ACM Trans. Netw. 12 (4) (2009) 1118–1131.

[22] A. Jindal, K. Psounis, On the efficiency of CSMA-CA scheduling in wireless multihop networks, IEEE/ACM Trans. Netw. 21 (5) (2013) 1392–1406.

[23] The Network Simulator NS-2. <http://www.isi.edu/nsnam/ns/>.

[24] P. Gupta, P.R. Kumar, The capacity of wireless networks, IEEE Trans. Inform. Theory 46 (2) (2000) 388–404.

[25] I. Rhee, A. Warrier, M. Aia, J. Min, Z-MAC: a hybrid MAC for wireless sensor networks, in: 3rd International Conference on Embedded Networked Sensor Systems, SenSys, 2005, pp. 90–101.

[26] G. Anastasi, M. Conti, M. Di Francesco, A. Passarella, Energy conservation in wireless sensor networks: a survey, Ad Hoc Netw. 7 (3) (2009) 537–568.

Ertugrul Necdet Ciftcioglu received his B.S. degree in Electrical and Electronics Engineering from the Middle East Technical University (METU), Ankara, Turkey in 2004, M.S. degree in Electronics Engineering and Computer Science from Sabanci University, Istanbul, Turkey in 2006, and Ph.D. degree in Electrical Engineering from The Pennsylvania State University, PA in 2012. He has been a Research Associate in the Department of Computer Science and Engineering, The Pennsylvania State University from 2012 to 2014. He has also been a visiting scholar at Northwestern University, IL, University of Southern California, CA and BBN Technologies, MA. Since September 2014, he is a Postdoctoral Researcher at IBM Research, Yorktown Heights, NY. His research interests are cross-layer design and resource allocation for wireless communication networks, particularly stochastic network optimization for relaying, cooperative communications and multiuser networks, and recent emphasis on network science.

Antonios Michaloliakos graduated in 2009 from the department of Electrical and Computer Engineering of the National Technical University of Athens. He is currently pursuing a PhD in Electrical Engineering at the University of Southern California. His interests span the areas of probabilistic modeling and design of networks and cross-layer optimization in new generation wireless networks. He is currently working on analytic modeling of MU-MIMO systems.

Aylin Yener received the B.Sc. degree in electrical and electronics engineering, and the B.Sc. degree in physics, from Bogazici University, Istanbul, Turkey; and the M.S. and Ph.D. degrees in electrical and computer engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, New Brunswick, NJ. Commencing fall 2000, for three semesters, she was a P.C. Rossin Assistant Professor at the Electrical Engineering and Computer Science Department, Lehigh University, PA. In 2002, she joined the faculty of The Pennsylvania State University, University Park, PA, where she was an Assistant Professor, then Associate Professor, and is currently Professor of Electrical Engineering since 2010. During the academic year 2008–2009, she was a Visiting Associate Professor with the Department of Electrical Engineering, Stanford University, CA. Her research interests are in information theory, communication theory and network science, with recent emphasis on green communications and information security. She received the NSF CAREER award in 2003. Dr. Yener previously served as a technical program chair or co-chair for various conferences for the IEEE Communications Society, as an associate editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, as an associate editor and an editorial advisory board member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. She served as the student committee chair for the IEEE Information Theory Society 2007–2011, and was the co-founder of the Annual School of Information Theory in North America co-organizing the school in 2008, 2009 and 2010. Dr. Yener currently serves on the board of governors of the IEEE Information Theory Society as its treasurer.

Konstantinos Psounis is an associate professor of Electrical Engineering and Computer Science at the University of Southern California. He received his first degree from the department of Electrical and Computer Engineering of National Technical University of Athens in 1997, the M.S. degree in Electrical Engineering from Stanford University, California, in 1999, and the Ph.D. degree in Electrical Engineering from Stanford University in 2002. Konstantinos models and analyzes the performance of a variety of wired and wireless networks and designs schemes and protocols to solve problems related to such systems. He is the author of numerous research papers on these topics which have received thousands of citations, and has faculty awards from multiple sources including the National Science Foundation, the Army Research Laboratory and CISCO Systems. He is a senior member of both IEEE and ACM.

Thomas F. La Porta is the William E. Leonhard Chair Professor in the Computer Science and Engineering Department at Penn State. He received his B.S.E.E. and M.S.E.E. degrees from The Cooper Union, New York, NY, and his Ph.D. degree in Electrical Engineering from Columbia University, New York, NY. He joined Penn State in 2002. He is the Director of the Institute of Networking and Security Research at Penn State. Prior to joining Penn State, Dr. La Porta was with Bell Laboratories since 1986 where he was the Director of the Mobile Networking Research Department in Bell Laboratories, Lucent Technologies where he led various projects in wireless and mobile networking. He is an IEEE Fellow, Bell Labs Fellow, received the Bell Labs Distinguished Technical Staff Award in 1996, and an Eta Kappa Nu Outstanding Young Electrical Engineer Award in 1996. He also won a Thomas Alva Edison Patent Awards in 2005 and 2009. His research interests include mobility management, signaling and control for wireless networks, security for wireless systems, mobile data systems, and protocol design. Dr. La Porta was the founding Editor-in-Chief of the IEEE Transactions on Mobile Computing and served as Editor-in-Chief of IEEE Personal Communications Magazine.

Ramesh Govindan received his B. Tech. degree from the Indian Institute of Technology at Madras, and his M.S. and Ph.D. degrees from the University of California at Berkeley. He is a Professor in the Computer Science Department at the University of Southern California. His research interests include routing and measurements in large internets, wireless sensor networks, and mobile computing systems.