



HAL
open science

Game Theoretical Analysis of the Tradeoff Between QoE and QoS Over Satellite Channels

Smrati Gupta, Elena Veronica Belmega, Maria Ángeles Vázquez-Castro

► **To cite this version:**

Smrati Gupta, Elena Veronica Belmega, Maria Ángeles Vázquez-Castro. Game Theoretical Analysis of the Tradeoff Between QoE and QoS Over Satellite Channels. 7th Advanced Satellite Multimedia Systems Conference/13th Signal Processing for Space Communications Workshop (ASMS/SPSC 2014), Sep 2014, Livorno, Italy. hal-01308853

HAL Id: hal-01308853

<https://hal.science/hal-01308853>

Submitted on 28 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Game Theoretical Analysis of the Tradeoff Between QoE and QoS Over Satellite Channels

Smrati Gupta

Dept. of Telecom. & Systems Engineering
Univ. Autonoma de Barcelona,
Spain

E. Veronica Belmega

ETIS UMR 851
ENSEA-Univ. Cergy-Pontoise-CNRS,
France

M. A. Vázquez-Castro

Dept. of Telecom. & Systems Engineering
Univ. Autonoma de Barcelona,
Spain

Abstract—In this work, a game-theoretic analysis of a video exchange application in which two users exchange their video streams over a satellite channel using Quality of Experience (QoE) driven rate adaptation is studied. In such an interaction, users aim at maximizing the Quality of Service (QoS) and QoE of their received video while minimizing their individual cost incurred by their video transmission, which is modeled as a repeated game. Given the payoff model of the users, it is shown that adaptive video exchange between selfish autonomous nodes for a deterministic time will not be sustained. However, if video is exchanged over an unlimited or indeterminate period, the nodes have an incentive to cooperate and exchange video streams with QoE-driven rate adaptation based on the trust they build among themselves. Our simulation results show that a tradeoff exists between the QoS and the QoE of the perceived video. Furthermore, it is shown that the expected interaction length has a high impact on such tradeoff.

I. INTRODUCTION

With the recent growth in demands of satellite services like video applications based services, there has been a lot of research to find mechanisms for efficient use of limited resources in satellite communications. Such mechanisms include the design of transmission rate adaptation schemes to utilize and adapt to the dynamic satellite environment [1], [2]. These rate adaptation schemes are designed on the basis of various factors like satisfaction of the end users' demands, optimization of resource utilization etc.. However, due to the limited and expensive resources in satellite communications, there is competition among the terrestrial user nodes over existing resources, and therefore an important factor influencing the performance of rate adaptation schemes is the competitive behavior among these nodes. It is needed to design the schemes with an in-depth analysis of such behavior.

Game theoretical framework allows us to capture this competitive behavior under the light of natural satellite environment limitations. The use of game theoretical framework to study the problems in satellite networks has been mostly focussed on resource allocation among the nodes (e.g., [3], [4]). However, the game theoretical analysis of interaction between two nodes employing complex rate adaptation strategies has been quite limited and focussed upon single shot games [5], [6]. Therefore, in this paper, we study a video exchange application over the satellite channels among two nodes using game theoretical tools, in particular, under the light of repeated games framework. Our contributions can be summarized as follows:

First, with the help of repeated games' framework, we identify the best performing QoE-driven rate adaptation strategies for adaptive video transmission in different satellite networks. We show that our framework is similar to the repeated prisoners dilemma [7] and that the adaptive video exchange between selfish autonomous nodes for a limited time will not be sustained. However, if video is exchanged over an unlimited period of time, the nodes have an incentive to cooperate and exchange video stream with QoE-driven rate adaptation based on the trust they build among themselves. The details of the proofs are omitted because of the lack of space. They follow in the similar lines as [8] in which the authors studied the interaction among two agents engaged in a competitive privacy interaction in an information-theoretic distributed source coding scenario.

Second, with the help of numerical simulations, we illustrate that the speed of rate adaptation used by the nodes is influenced by two factors: the expected satellite channel conditions and the expected length of interaction between the terrestrial user nodes in the repeated game framework. Our simulations show that as the expected length of interaction increases, the speed of rate adaptation to the channel conditions increases thus providing better QoS. Good channel conditions require a smaller length of interaction to achieve faster adaptation rates as compared to bad channel conditions. It turns out that these two factors also imply a tradeoff between QoS and QoE. This tradeoff is analyzed in details via numerical results.

This paper is organized as follows: in section II, we describe the system model and QoE-driven rate adaptation model. In section III, we describe the game theoretical approach applied to the adaptive video exchange between the users. The numerical simulations and analysis is presented in section IV, followed by the conclusions in section V.

II. SYSTEM MODEL AND QOE-DRIVEN RATE ADAPTATION

In this section, we describe the system model and the QoE-driven rate adaptation scheme used. We also define the performance metrics which will be considered in the analysis further.

A. System Model

1) *Network Topology*: Our focus is on a scenario in which two terrestrial nodes want to exchange their videos via a satellite relay. Such scenarios model situations like video

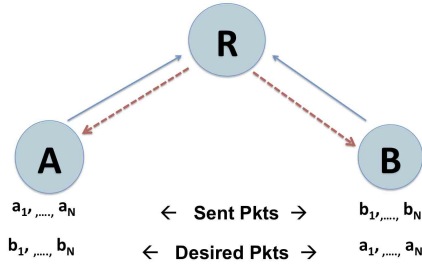


Fig. 1: Network Topology. The nodes A and B exchange information via R. The arrows indicate the links with positive channel capacity.

streaming between two parties across geographically separated areas. The node A (or B) transmits the video packets to the satellite and the satellite forwards these packets to node B (or A) as shown in figure 1. The wireless channel between any two nodes has variable capacity to transmit packets across the two nodes depending on various factors like environmental conditions (rain, cloud, etc.), congestion due to other users, etc. We model channel links between the nodes to have a certain throughput capacity which varies with time. The throughput capacity is the maximum number of packets that can be transmitted between any node k and node l at time t and is denoted by $R_{kl}(t)$. It is measured in packets per time slot (pps). As shown in figure 1, $R_{AB} = 0$ for all t . Throughout this paper, for simplicity, at any time instant t , all links are assumed to have equal throughput capacities $R_c(t)$, that is, $R_{kl}(t) = R_c(t), \forall kl \in \{AR, BR, RA, RB\}$. Each time instant is described by $t = n\Delta$, where $n \in \mathbb{Z}^+$ and Δ is the time interval size, i.e., a precision parameter. We will therefore, refer to instantaneous quantities using the integer value n . We also assume that the satellite node simply forwards the packet without any on-board processing. Therefore, $R_c(n)$ is the instantaneous end to end throughput capacity of the satellite channel between A and B and vice versa.

2) *Channel Model*: We model the satellite channel using a 2-state Markov channel model as described in [9]. The two-state markov chain model is shown in figure 2. The 'good' channel state corresponds to the channel being in good conditions, like line of sight situations, clear weather s etc.. The 'bad' channel state refers to channel being in bad conditions for example in situations when there is obstruction to transmission across the channel like clouds, rains etc. The markov chain model in figure 2 shows the state diagram for transition from one state to another. The probability of the channel to move from good state to bad and vice versa is p_{GB} and p_{BG} respectively. Consequently, the probability to remain in good and bad state are $p_{GG} = 1 - p_{GB}$ and $p_{BB} = 1 - p_{BG}$. From this model, the probability of the channel to be in bad state p_B and probability of channel to be in good state p_G is given by

$$p_B = \frac{p_{GB}}{p_{GB} + p_{BG}} \quad (1)$$

$$p_G = \frac{p_{BG}}{p_{GB} + p_{BG}} \quad (2)$$

We define that in good state the channel has a throughput capacity of R_1 whereas in bad state it has R_2 , such that $R_1 >$

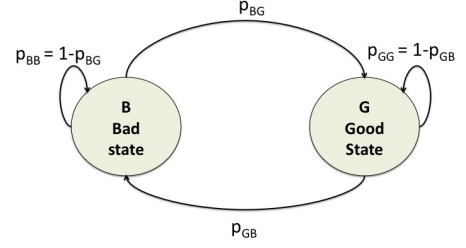


Fig. 2: Two State Markov Model for Satellite Channel

R_2 . Therefore, at any time instant $t = n\Delta$,

$$R_c(n) = \begin{cases} R_1 & \text{good state} \\ R_2 & \text{bad state} \end{cases} \quad (3)$$

In this paper, we assume that the channel coherence time is given by $T_c = c\Delta$ slots where $c \in \mathbb{Z}^+$. This means that given the channel is in a state, the channel state will not vary for at least $c\Delta$ slots. This assumption is reasonable as the satellite channel conditions, which are influenced by the environment, usually take time to vary. Another assumption is that each video stream exchange lasts for $T = N\Delta$ slots and $T \gg T_c$. Note that for better accuracy of channel model, this analysis can also be extended for the three-state Markov channel model [10].

3) *Cross-layer Design Model*: The video content is transmitted using UDP (User Datagram Protocol) over RTP (Real-Time Transport Protocol). The source node is equipped with a codec responsible for compression of the video content. The output rate of the codec can be reconfigured to deliver a target bit-rate. Therefore, at the source node, the output rate can be adjusted as desired. This output rate from any node $k \in \{A, B\}$ at time $t = n\Delta$ is denoted by $r_k(n)$ and it is measured in packets per time slot. The packets are further passed down to the network layer maintaining coherence with standard protocol stack. The rate of transmission of video payload is adapted to the network conditions by optimization of the QoE perceived at the end-user. With the use of Real Time Control Protocol (RTCP) signaling, the source node collects the feedback information about the Round Trip Time (RTT) from the destination. The RTCP feedback signaling provides the source node with information to re-configure the codec rate to suit the target needs.

4) *Buffer and Delay Model*: Each of the source nodes (A and B) maintains a transmission buffer of size B_c packets. The number of packets stored in transmission buffer at time $t = n\Delta$ is given by $B(n)$. The buffer is maintained using a First In First Out (FIFO) queue. Let the source node k transmit the packets at time $t = n\Delta$ at the rate $r_k(n)$. The source node is unaware of the channel link capacity $R_c(n)$. Assume that the buffer is initially empty. If $r_k(n) < R_c(n)$, all the packets are transmitted through the channel with rate $r_k(n)$. If $r_k(n) > R_c(n)$, then packets are transmitted with a rate of $R_c(n)$ through the channel link. From the remaining $r_k(n) - R_c(n)$ packets per time slot, $B(n) \leq B_c$ packets are stored in the empty space in the transmission buffer, while the remaining

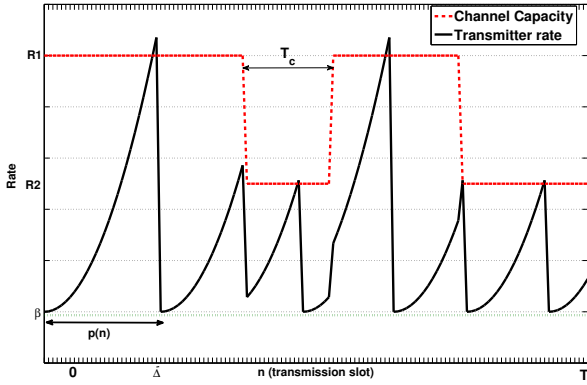


Fig. 3: Illustration of Rate adaptation curve. The red line indicates the rate $r_k(n)$. The blue line indicates the channel capacity. Angle α remains constant for all time n , whereas period $p(n)$ depends on channel capacity. It is assumed that $T \gg T_c$.

packets are lost. In the next time slot, the packets stored in the buffer are transmitted through the channel. These packets face a delay in reaching the destination. The same process as above is applied to rest of the new packets to be transmitted, but now the effective channel capacity is $R_c(n) - B(n-1)$.

Using the above system model, we will now describe the Quality of Experience-driven rate adaptation model which is used in the rest of the paper. Furthermore, we describe the pertinent performance metrics derived using [11] which take into account the applicability of the system model.

B. QoE-driven Rate Adaptation and Performance Metrics

The cross layer rate adaptation model shows the framework to optimize the output codec rate of a source node with feedback from the network in the form of RTCP signals. We will focus on the rate adaptation based on optimizing Quality of Experience of the video obtained by the end-user. The QoE of the end user can be quantified using different metrics in spatial domain like SSIM (structural symmetry) and temporal domain like flow continuity. In this work, we focus on using flow continuity of the end-user video as a measure of the QoE in the time domain. This is done in order to ensure that the end-user is able to view the video without any freezing. The rate is adapted to the channel conditions as follows. The source begins the transmission at a certain initial rate. The rate is increased with time, as long as the source observes no delay. However, as soon as the source observes a delay, it reduces the rate to its initial value. The process is again repeated. This adaptation is shown in figure 3. Mathematically, any source k begins the transmission at rate $r_k(1) = \beta < R_c(1)$, where $\beta \in \mathbb{R}^+$. With each RTCP received, the rate is updated by

$$r_k(n, \alpha_k) = \begin{cases} \beta + \delta r_k(n, \alpha_k) & \text{if no delay observed} \\ \beta & \text{if delay observed} \end{cases} \quad (4)$$

where the increment in rate is

$$\delta r_k(n, \alpha_k) = [\alpha_k(n \bmod p(n))\Delta]^2$$

such that Δ is the time interval size, $\alpha_k \in (0, \tan(\pi/2))$ is the slope of the rate adaptation curve and $p(n)$ is the time period of the waveform in figure 3 given by

$$p(n) = \left\lceil \frac{\sqrt{R_c(n) - \beta}}{\alpha_k \Delta} \right\rceil + 1$$

Note that this rate adaptation is done at the source node k , and it is adapted according to the delay statistics observed at node $l \neq k$. Henceforth, we denote any node other than k as $-k$. Note that the increment in rate can be modeled using different functions like logarithmic function, linear function or exponential function etc. In this paper, we have assumed an squared function to model the rate increment having a slow start followed by a rapid growth. Such a function models rate adaptation in different protocols like TCP variants which are slow start [12], [2].

We now describe the performance metrics which will be used to analyze the performance of the scheme.

1) *Quality of Service Metric*: The network utility is defined as the ratio of the network capacity utilized to transmit packets to a node and the total network capacity provided by the network. The number of packets obtained at node k depends on the rate of transmission at the other node $r_{-k}(n, \alpha_{-k})$. However, the packets obtained at the node k cannot exceed the instantaneous channel capacity. Therefore, the instantaneous network utility at time $t = n\Delta$ is given by

$$\mu(R_c, r_{-k}, n, \alpha_{-k}) = \frac{\min(R_c(n), r_{-k}(n, \alpha_{-k}))}{R_c(n)}$$

where $R_c(n)$ is given by (3) and $r_{-k}(n, \alpha_{-k})$ is given by (4). We define the averaged network utility from $n = 0$ to $n = N$ in providing the video to node k as the *Quality of Service* at node k and it is given by

$$f_k^{QoS}(\alpha_{-k}) = \frac{1}{N} \left\{ \sum_{n=0}^N \mu(R_c(n), r_{-k}(n, \alpha_{-k})) \right\} \quad (5)$$

It can be seen that the QoS of node k depends on the rate adapted by the other node and consequently on α_{-k} .

2) *Quality of Experience Metric*: The flow continuity is defined as the probability of the delay in the network not to exceed the threshold delay. The delay in the network leads to video packets arriving at the end user in more than expected time, leading to a visible freezing of the displayed video and degrading the quality of experience. In this paper, flow continuity is chosen as the QoE metric which is a simplified choice for the analysis. Assuming the threshold delay to be 0 (allowing no freezing of video), we notice in the buffer and delay model described in section II-A4 that the delay is observed at the node k when the rate r_{-k} exceeds the channel capacity. Let us define a delay counter function $\varphi(\cdot)$ which determines if there is an instantaneous delay in the network or not. It takes the instantaneous rate and the channel capacity as the input and generates 0 if there is no delay or 1 if there is a delay in the network. It is defined as

$$\varphi(r(n, \alpha), R) = \max(0, \text{sgn}(r(n, \alpha) - R))$$

where

$$\text{sgn}(r(n, \alpha) - R) = \begin{cases} -1, & \text{if } r(n, \alpha) < R \\ 0, & \text{if } r(n, \alpha) = R \\ 1, & \text{if } r(n, \alpha) > R \end{cases}$$

Therefore, we can write the total number of events when the delay occurs as

$$\sum_{n=0}^N \varphi((r_{-k}(n, \alpha_{-k}), R_c(n)))$$

We define the *Quality of Experience* metric as the average flow continuity of the network and it is given by

$$f_k^{QoE}(\alpha_{-k}) = 1 - \left\{ \frac{\sum_{n=0}^N \varphi(R_c(n), r_{-k}(n, \alpha_{-k}))}{N} \right\} \quad (6)$$

We note that the flow continuity at k is a function of the rate adaptation gradient at the other node α_{-k} .

3) *Cost*: In order to transmit packets, the sender node k incurs a cost of transmission due to the power usage, hardware requirements etc. This cost is dependent on not only the magnitude of resources used for transmission but also on the gradient of increment in the rate. This is because the higher the gradient in rate increment, the more is the difference between two consecutive instantaneous rates. This asserts a higher energy demand on the resources to make a sharper change in the rate leading to higher cost. In order to model this cost of transmission, for simplicity, we use a logarithmic function of the rate adaptation slope or α given by

$$f_k^{COST}(\alpha_k) = \log_2(1 + \arctan(\alpha_k)) \quad (7)$$

Note here that the cost of transmission incurred at the node k is determined by the rate adaptation gradient used by the node k .

With these metrics, in the next section, we will define the utility function given by

$$w_1 f_k^{QoS}(\alpha_{-k}) + w_2 f_k^{QoE}(\alpha_{-k}) - w_3 f_k^{COST}(\alpha_k)$$

where $w_i \in \mathbb{Z}$ and it will be optimized further in the game-theoretical analysis.

III. GAME THEORETICAL ANALYSIS FOR ADAPTIVE SATELLITE VIDEO TRANSMISSION

It can be seen from section II-A that the quality of the video obtained by a node, say A , depends on the rate of transmission of node B and vice versa. Additionally, in order to transmit the video packets to node A , node B incurs a cost and vice versa. Therefore, for two selfish nodes A and B , there is a conflict of interest as both nodes want to incur minimum cost (affecting the video quality of other user) and also obtain good quality of video themselves (affected by the rate of transmission by other user). Therefore, there is an interaction arising naturally among the two nodes which is modeled using game theory.

We assume that each of this interaction occurs for a video stream exchange lasting for $T = N\Delta$ slots, which means that each video stream exchange session lasts for at least T slots. We will now study the model of the game when the

two nodes exchange video stream only once (lasting for T slots). In the subsequent sections, we will consider how this interaction varies if the nodes exchange the video streams repeatedly where each exchange lasts for T slots.

A. One-Shot Game Analysis

We define a one-shot non-cooperative game by the following tuple:

$$\mathcal{G}_O = \{\mathcal{P}, \{\mathcal{A}_k\}_{k \in \mathcal{P}}, \{u_k\}_{k \in \mathcal{P}}\} \quad (8)$$

where \mathcal{P} refers to the set of players which selfishly maximize their own payoffs given by u_k by choosing their actions from the action set \mathcal{A}_k for every $k \in \mathcal{P}$. With the QoE-driven rate adaptation model to exchange video, we define the following component sets of the game tuple \mathcal{G}_O :

Players Set \mathcal{P} : The set of players is given by $\mathcal{P} = \{A, B\}$ which are the two nodes A and B that exchange the video packets.

Action Set \mathcal{A}_k : The set of actions that can be taken by the player k , where $k \in \{A, B\}$, is given by $\mathcal{A}_k = [0, \lambda \frac{\pi}{2}]$ and $\lambda \in (0, 1)$. The action chosen by k -th player from \mathcal{A}_k is denoted by α_k . Hence the player can choose the gradient of rate adaptation as its action in order to maximize its own benefits. The factor λ ensures that the gradient is always $\alpha < \frac{\pi}{2}$ to prevent infinite increase in rate.

Utility u_k : The utility, or payoff function of the k -th user is defined as follows

$$u_k(\alpha_k, \alpha_{-k}) = w_1 f_k^{QoS}(\alpha_{-k}) + w_2 f_k^{QoE}(\alpha_{-k}) - w_3 f_k^{COST}(\alpha_k) \quad (9)$$

The utility of the k -th player is a function of the action α_k taken by the node k and the action α_{-k} taken by the other node. The components of utility function are as follows: the utility is a joint measure of the quality of video experienced by the node k and the cost incurred in the transmission of video to the other node. The video quality at node k is affected by the throughput and the flow-continuity given by $f_k^{QoS}(\alpha_{-k})$ and $f_k^{QoE}(\alpha_{-k})$, using (5) and (6) respectively, which are both determined by the action of the other node α_{-k} . Since the node k wants to maximize its quality of video perceived, the utility is directly proportional to both throughput and flow continuity. However, the node k incurs a cost $f_k^{COST}(\alpha_k)$, given by (7), to transmit video to the other node, which is inversely proportional to its utility. Finally the weight $w_i \in (0, 1)$ assigns the appropriate dimensions to different factors such that $\sum_{i=1}^3 w_i = 1$.

Consider the scenario when the two players interact only once for T slots to exchange the video streams. This leads to the one shot game \mathcal{G}_O in (8). Before the players begin to exchange, they decide their actions, i.e. at which gradient they will transmit during the next T slots. The Nash Equilibrium (NE) of a non-cooperative game \mathcal{G} is defined as a set of actions of the players $(\alpha_k^*, \alpha_{-k}^*)$, from which no player has an incentive to deviate unilaterally. Hence, the selfish rational players are expected to play the action at NE irrespective of the other players action.

In the one-shot game, when the players cannot build trust with each other, they selfishly choose the action to maximize their utility. Node k can maximize its utility by minimizing the cost given by (7) at $\alpha_k = 0$. The node k has no control over the action taken by the other user, therefore, the part of its utility affected by the other user's action, i.e., its QoS and QoE, cannot be controlled. Given that the strictly dominant strategy for each user is not to cooperate, the following result follows [7]:

Lemma 1. *The unique Nash Equilibrium of one-shot QoE-driven adaptive video exchange game $\mathcal{G}_O = \{\mathcal{P}, \{\mathcal{A}_k\}_{k \in \mathcal{P}}, \{u_k\}_{k \in \mathcal{P}}\}$ is given by $(\alpha_1^*, \alpha_2^*) = (0, 0)$.*

The intuition is that no selfish end user would be willing to incur a cost to send the video when there is no guarantee of receiving any video in return.

B. Repeated Game Analysis

We will now present how the results of the interaction between the two players vary when the video exchange (of T slots each) takes place repeatedly. A repeated game, in which the game \mathcal{G}_O defined in (8) is played repeatedly, is defined using the following tuple:

$$\mathcal{G}_R = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, T_R\} \quad (10)$$

The set \mathcal{P} refers to the set of players, given by the set of nodes $\{A, B\}$.

The set \mathcal{S}_k is the strategy set of user k . The strategy set is different from the action set in (8) because it describes the actions taken by the player k based on the history of the play and on the actions at each instant. More precisely, let the actions taken by the players in the τ -th stage be $a^{(\tau)} = (a_1^{(\tau)}, a_2^{(\tau)})$. Then the history of the game at the end of stage $t \geq 1$ is given by $h^{(t+1)} = (a^{(1)} a^{(2)} \dots a^{(t)})$. The set of all possible histories up to t is given by $\mathcal{H}^{(t)} = \{\mathcal{A}_k \times \mathcal{A}_{-k}\}^t$ where \mathcal{A}_k is the action set in (8) given by $[0, \lambda \frac{\pi}{2}]$. The strategy of a player k is $s_k = (s_k^{(1)}, s_k^{(2)}, \dots, s_k^{(T)}) \in \mathcal{S}_k$ and maps each possible history to an action, as $s_k^{(t)} : \mathcal{H}^{(t)} \rightarrow \mathcal{A}_k$ such that $s_k^{(t)}(h^{(t)}) = a_k^{(t)}$.

The long term utility for the repeated game is a discounted utility function v_k such that for a strategy profile $s = (s_1, s_2)$, is the weighted sum of the stage utilities obtained by user k following the strategy s , such that the user discounts the future payoffs by a discount factor δ . It is given by

$$v_k(s) = \frac{(1 - \delta)}{(1 - \delta^T)} \sum_{t=1}^T \delta^{t-1} u_k(a^{(t)}) \quad (11)$$

where $a^{(t)}$ is the action profile at stage t induced by strategy s (i.e. $s_k^{(t)}(h^{(t)}) = a_k^{(t)} \forall k$), discount factor $\delta \in (0, 1)$ and u_k is the one-stage payoff function defined in (9).

The parameter $T_R > 1$, refers to the number of times the interaction takes place.

Before we discuss the outcome of the repeated game \mathcal{G}_R , we briefly describe the concept of sub-game perfect equilibrium for repeated games [13]. In coherence with the NE of one-shot game, the NE of repeated game is a strategy profile

from which no player gains by unilaterally deviating. However, there are some strategy profiles, which are not expected to occur due to rational behavior of the players. Hence the NE region can be narrowed down to define a subgame perfect Equilibrium region, as a subset of NE. A subgame is a game from stage t onwards having a history $h^{(t)}$, and is denoted by $\mathcal{G}_R(h^{(t)})$. The strategies and payoffs used for the sub-game are the functions of possible histories that are consistent with $h^{(t)}$. A subgame perfect equilibrium (SPE) is defined as a strategy profile $s^* = (s_1, s_2)$ such that for all $h^{(t)} \in \mathcal{H}^{(t)}$, the strategy $s^* | h^{(t)}$ is a NE for any subgame $\mathcal{G}_R(h^{(t)})$.

Consider the scenario when the two players begin their video exchange. Both players are aware of the fact that the transmission is going to be repeated for $T_R > 1$ and hence they play the game \mathcal{G}_R in (10). Before the players begin the transmission, they decide their strategy i.e., at what gradient they both should adapt their rates. This decision depends on the perception of the players of whether T_R is finite or infinite.

1) Finite Horizon Repeated Game: In this case, the two players exchange the video streams with each other (each exchange lasting for T slots) repeatedly but a finite number of times given by $T_R < \infty$. In such a case, the players are aware that they will interact only for a limited time. Similarly to the repeated prisoners' dilemma, let us consider the last stage of the game, when it is played T_R -th time. The players are aware that they interact for the last time. They have no incentive to transmit the video for the other player while incurring a cost themselves, when there is no guarantee that the other player will transmit the video or not. Hence, their optimal strategy is $s_k^{(T_R),*} = 0$. Now, when the players play the game at time $T_R - 1$, given that in stage T_R they will not transmit, there is no incentive to transmit at stage $T_R - 1$. Therefore, $s_k^{(T_R-1),*} = 0$. Using this principle, the following lemma can be deduced:

Lemma 2. *The unique sub-game perfect equilibrium of the video exchange game between two nodes, defined by $\mathcal{G}_R = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, T\}$, where $T < \infty$ and is known to both nodes is given by*

$$s_k^{(t),*} = 0, \quad \forall k \in \mathcal{P}, \forall t \in \{1, 2 \dots T\}$$

Given that the players are aware of the number of times the video exchange will occur, they act selfishly at each stage of the game in order to maximize their payoffs. There is no incentive in building trust in this case.

2) Infinite Horizon Repeated Game: In this case, the players will interact and exchange the video streams (lasting for T slots), but such interaction lasts for an infinite time T_R . In other words, the players are not aware of how many times they will exchange the video with the other user. The reason why cooperation (i.e., any strategy apart from $(0, 0)$) is not sustainable in finite horizon games is that the players know precisely when the interaction ends. However, in infinite horizon repeated games, the players in the game \mathcal{G}_R do not know precisely when the game will end, or equivalently, $T_R \rightarrow +\infty$. We will now describe some of the achievable SPE for such games. Note that the overall achievable SPE region for the infinite horizon repeated games is an open problem and not known in general [7], [13]. We will consider two of such achievable SPE.

In an infinite horizon repeated game, when the players begin to interact, they decide their strategy based on the information that the interaction will last for an infinite amount of time. One possible SPE for such games is given by the non-cooperative strategy at each stage of the game (give by $(0, 0)$). We will focus here on more interesting SPE in which non-trivial cooperation can be achieved. The intuition for existence of other SPE is similar to the infinite time horizon prisoner's dilemma [7]. At any stage of the game, the pay-off in (9) has two components: (i) the component controlled by the opposite player's action which increases the utility, but the player k has no control over it (ii) the component influenced by the player's own action which decreases the utility. The player has no control over the first component, but in long term, it can build trust with the other player to influence the other player's decision and thereby both players can obtain a long term sustainable utility which is higher than utility achieved by the $(0, 0)$. We focus on such an action profile to evaluate the long term sustainability of the action profile whose utility is better than the one-shot NE for both players and are given by

$$\begin{aligned} u_1(\alpha_1^*, \alpha_2^*) &> u_1(0, 0) \\ u_2(\alpha_1^*, \alpha_2^*) &> u_2(0, 0) \end{aligned} \quad (12)$$

In the following lemma, we identify the SPE which can achieve non-trivial outcomes due to the trust built among the players.

Lemma 3. *In an infinite-horizon repeated game $\mathcal{G}_{\mathcal{R}} = \{\mathcal{P}, \{\mathcal{S}_k\}_{k \in \mathcal{P}}, \{v_k\}_{k \in \mathcal{P}}, +\infty\}$, and with an agreement profile (α_1^*, α_2^*) satisfying (12), if the discount factor is bounded by*

$$1 > \delta > \delta^{\min}(\alpha_1^*, \alpha_2^*) \quad (13)$$

where $\delta^{\min}(\alpha_1^*, \alpha_2^*)$ is given by (14) on the next page, then the following strategy is an SPE: "A node k transmits with gradient α_k^* at the first stage and continues to adapt the rate with this gradient as long as the other player adapts its rate at least by α_{-k}^* . If any player has ever defected, then the players transmit at $\alpha = 0$ for the rest of the interaction."

The proof of the above can be derived in a similar way as in [8] using the one-step deviation principle. The discount factor can be seen as the perception of the players of the probability for the game not to end at every stage of the game. If the players perceive the probability of the game not to stop above certain limits, there can be a non-trivial SPE. In other words, if the probability of the game not to stop is large enough, the players would rather prefer to develop trust and obtain better payoffs than minimize their one-shot costs. We also note here, that only those points which satisfy (12) are sustainable thereby ensuring that the players agree upon the points which offer them incentive over one-shot NE.

IV. NUMERICAL RESULTS AND ANALYSIS

We will now present simulation results to identify how the satellite channel conditions modify the outcome of the infinite horizon repeated game. We assume that if the channel is in good state then $R_1 = 350$ kbps and if the channel is in bad state, then $R_2 = 200$ kbps. The channel coherence time is $T_c = 20s$ whereas the length of time for which each video exchange lasts is $T = 20000s$. The other variables are fixed as $\beta = 150$ kbps, $\Delta = 1$, $w_1 = w_2 = 0.45$ and $w_3 = 0.1$.

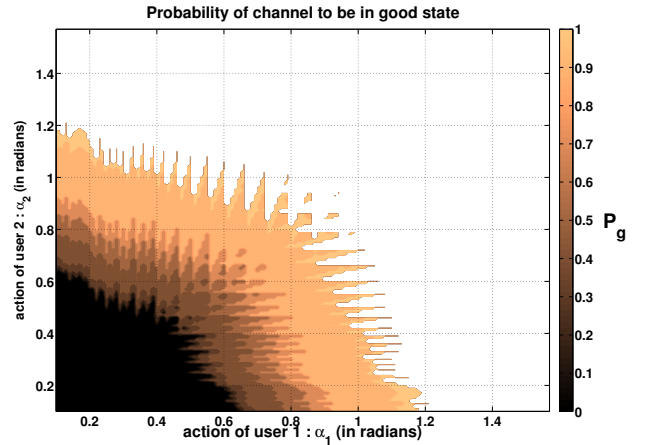


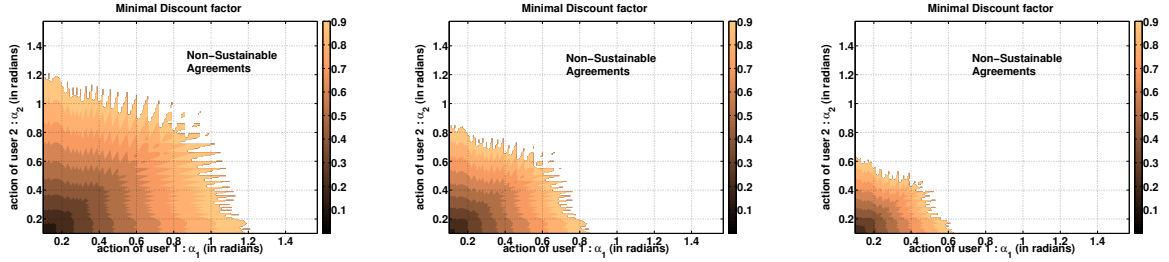
Fig. 4: Achievable region increases as channel conditions improve and there is a higher probability of channel to be in good conditions

A. Effect of Channel Conditions on Game Outcome

We consider the scenario where two nodes exchange video repeatedly for an undefined horizon of time. We identify the action profiles (rate gradients), apart from $(0, 0)$, which the players agree to transmit at given in Lemma 3. Given our varying channel model, the probability of channel to be in good state (p_g) also affects the achievable agreement points. In figure 4, we plot the achievable agreement region based on varying channel conditions. We assume $\delta \rightarrow 1$ which implies the probability of game to stop approaches 0. When p_g tends to 1, there is a high probability that channel is in a good condition. This leads to a higher probability that payoff obtained by the players at action profiles farther away from $(0, 0)$ exceeds payoff obtained at $(0, 0)$ and such action profiles are achievable based on channel conditions (12). However, when channel conditions are bad and p_g tends to 0, the magnitude of payoff at action profiles farther away from $(0, 0)$ is not high enough to exceed payoff at $(0, 0)$. Hence, a smaller agreement region is achievable in bad channel conditions than in good channel conditions. This implies that the players will adapt to the channel conditions faster, with a higher gradient, when the conditions are good, thereby utilizing the available resources more efficiently. However, in bad channel conditions, the players will not agree to faster adaptation to the channel conditions. Hence, bad channel conditions force the players to be cautious and the available resources remain under-utilized.

Also note, the curve has spikes periodically due to the non uniform variation of utility function. The utility function $u_k(\alpha_k, \alpha_{-k})$ varies uniformly with α_k because $\frac{du_k(\alpha_k, \alpha_{-k})}{d\alpha_k}$ is monotonously decreasing for increasing α_k . This implies that increasing the user's own gradient simply incurs more cost to the user and hence decreases its own payoff. However, the utility function does not vary monotonously with α_{-k} . As α_{-k} increases, the flow continuity decreases due to more instances of rate exceeding the channel capacity (explained in detail in next subsections). In addition, as α_{-k} increases, the quality of service increases for smaller values of α_{-k} , when there are not many instances of rate exceeding the channel

$$\delta^{min}(\alpha_1^*, \alpha_2^*) = \max_{k \in \mathcal{P}} \frac{w_3 [f_k^{COST}(\alpha_k^*) - f_k^{COST}(0)]}{w_1 [f_k^{QoS}(0) - f_k^{QoS}(\alpha_{-k}^*)] + w_2 [f_k^{QoE}(0) - f_k^{QoE}(\alpha_{-k}^*)]} \quad (14)$$



(a) Probability of channel in good state = 1 (b) Probability of channel in good state = 0.5 (c) Probability of channel in good state = 0

Fig. 5: Variation in minimal discount factor within achievable region for various channel conditions

capacity. However, when such instances increase at high values of α_{-k} , the quality of service decreases with increasing α_{-k} . Therefore, due to the culmination of these two effects, the agreement point region does not show monotonous behavior. In fact, the derivative of utility function with action profile of other user is a complex trigonometric function which is periodic in nature leading to the shape of the curve being similar to periodic function.

B. Effect of Channel Conditions on Minimal Discount Factor

Now, we will show how the minimal discount factor necessary to achieve any point varies based on different channel conditions. More precisely, we show how high should be the probability for game to continue for the players to agree on a specific gradient in a given channel condition. In figure 5, we plot the minimal discount factors necessary to achieve different action profiles within the achievable region which is based on channel conditions. As the achievable region decreases from good channel in figure 5a to bad channel in figure 5c, the discount factor necessary to sustain an action profile also varies proportionally. This fact can be observed more closely in figure 6 in which we consider that the players agree on the symmetric action profiles (or equal gradients). It can be seen (in figure 6) that, as the channel conditions improve and p_g varies from 0 to 1, a lower probability of game to continue is needed for any action profile to be achievable. This implies that even when the probability of game to continue is low, the players have a high incentive in transmitting at a higher gradient to achieve a high utility because the channel condition is good. It can also be observed that the players will transmit at a higher gradient of $\alpha = 0.7$ only in good channel conditions. For players to agree on a lower gradient like $\alpha = 0.4$, there should be a high probability of game to continue in bad channel conditions, but in good channel conditions, this gradient can be achieved even with low probability of game to continue. Therefore, the rate of adaptation to channel by the players depends on the channel conditions as well as their perception of the game to continue.

C. Achievable QoS and QoE Tradeoff in Satellite Channel

We will now show the variation in the observed QoS and QoE at the user level based on the satellite channel conditions

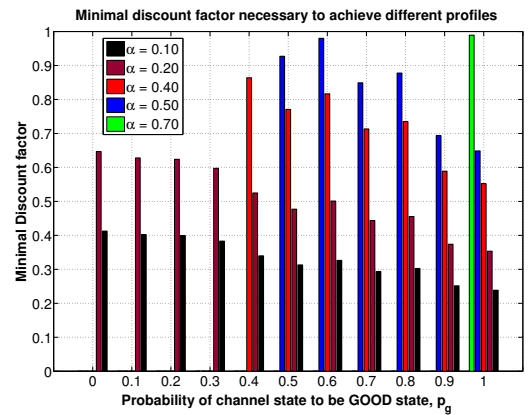


Fig. 6: Different minimal discount factor needed to achieve different symmetric action profiles for various channel conditions. The better is the channel, the lower is the discount factor needed to achieve profiles farther away from (0,0).

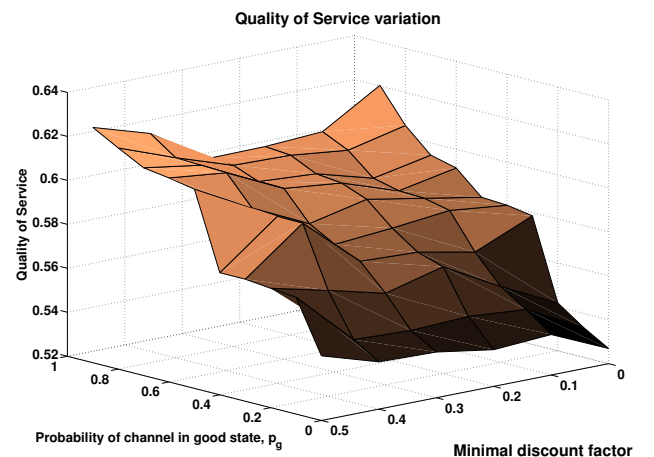


Fig. 7: QoS variation with varying discount factor and channel conditions

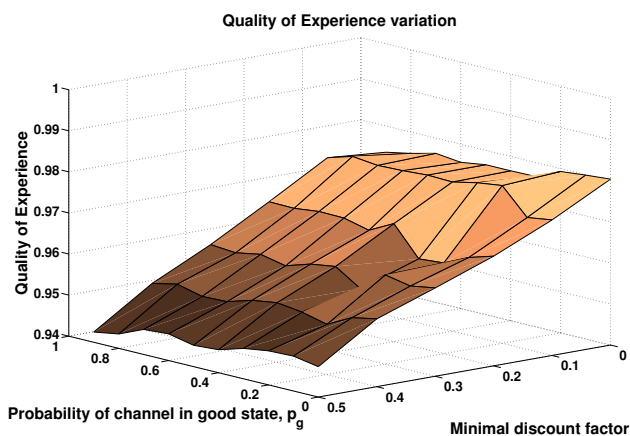


Fig. 8: QoE variation with varying discount factor and channel conditions

as well as the discount factor. Assuming that the players agree to only symmetric action profiles which are pareto optimal, in figure 7 and figure 8, we show the variation in QoS and QoE, observed when the game is played under varying channel conditions and minimal discount factor.

In figure 7, it can be seen that the overall QoS improves as the channel conditions improve. This is due to a higher number of packets transmitted when the channel has higher average throughput capacity. The channel conditions are not very crucial in obtaining the flow continuity or QoE, as seen in figure 8. This is because QoE depends on the probability of rate to exceed the channel capacity leading to delays. It is independent of the magnitude of the channel capacity.

Figure 7 also shows that as the discount factor increases, the QoS increases. However, in case of QoE, in figure 8, as the discount factor increases, the QoE decreases. As the discount factor increases, the probability of the game to continue also increases which allows the players to agree on higher rate adaptation gradients. This causes a higher number of packets to be transmitted per unit time which improves the overall QoS. However, at higher gradients, there is a higher chance of the rate to exceed the channel capacity, which leads to higher chances of delay and freezing of video resulting in lower QoE. Therefore, there exists a tradeoff between QoS and QoE depending on the length of the interaction between the players.

We now observe the joint effect of the two factors on QoS and QoE. As we saw in previous results, for a good channel higher adaptation gradients are achievable. This leads to higher QoS and lower QoE that are the extreme points in figure 7 and figure 8. However, in bad channel conditions lower adaptation gradients are achieved. This leads to lower QoS due to low throughput but higher QoE due to low probability of delay. Hence, depending on the perception of the user of the channel conditions and the length of interaction, the QoS and QoE are achievable. Overall, such effect of such factors can be used to potentially refine the rate adaptation algorithms.

V. CONCLUSIONS

In this paper, we have analyzed a QoE-driven adaptive video exchange between two selfish users (nodes) across a satellite channel and analyzed it using game theoretical tools. With the use of these tools, not only the selfish and realistic behavior of the users has been modeled but the expected outcomes of such an interaction have been predicted. It has been shown that the speed at which the users adapt the transmission rate to the channel conditions is influenced by the two fold effect of expected length of interaction between the users and the expected channel conditions. Consequently, the best strategies are chosen by the users based on available parameters which affects the QoS and QoE of the video achieved. It turns out that a tradeoff exists between the achieved QoS and QoE based on the expected length of interaction among the users. This work can be further explored to exploit more complicated satellite systems under the framework of repeated games.

REFERENCES

- [1] Sun, Jun, et al. "Ka-band link optimization with rate adaptation for Mars and lunar communications." *International Journal of Satellite Communications and Networking* 25.2 (2007): 147-165.
- [2] Pimentel-Niño, M. A.; Vázquez-Castro, M. A. & Skinnemoen, H. Optimized ASMIRA Advanced QoE Video Streaming for Mobile Satellite Communications Systems 30th AIAA International Communications Satellite Systems Conference, 2012.
- [3] Del Re, E.; Gorni, G.; Ronga, L.S.; Vázquez-Castro, M. A., "A Game Theory Approach for DVB-RCS Resource Allocation", Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE , vol., no., pp.2937,2941, 11-14 May 2008.
- [4] Vázquez-Castro, M. A.; Zhu Han; Hjørungnes, A.; Marina, N., "Rate allocation for satellite systems with correlated channels based on a Stackelberg game", *Game Theory for Networks*, 2009. GameNets '09. International Conference on , vol., no., pp.638,645, 13-15 May 2009.
- [5] Ye J, Hamdi M (2011) Priority-based rate adaptation using game theory in vehicular networks. In: *Global Telecommunications Conference (GLOBECOM 2011)*, 2011 IEEE, pp 1-6, DOI 10.1109/GLOBECOM.2011.6133998.
- [6] Trinh TA, Molnár S (2004) A game-theoretic analysis of tcp vegas. In: *Quality of Service in the Emerging Networking Panorama*, Springer, pp 338-347.
- [7] Fudenberg, D. & Tirole, J. *Game Theory* MIT Press, 1991.
- [8] Belmega, E.; Sankar, L. & Poor, H. Repeated games for privacy-aware distributed state estimation in interconnected networks *Network Games, Control and Optimization (NetGCoP)*, 2012 6th International Conference on, 2012, 64-68.
- [9] Lutz, E., "A Markov Model for Correlated Land Mobile Satellite Channels", *International Journal of Satellite Communications*, Wiley Subscription Services, Inc., A Wiley Company, 1996, 14, 333-339.
- [10] Vázquez-Castro, M. A.; Perez-Fontan, F., "LMS Markov model and its use for power control error impact analysis on system capacity", *Selected Areas in Communications*, IEEE Journal on , vol.20, no.6, pp.1258,1265, Aug 2002.
- [11] Gupta, S.; Pimentel-Niño, M. A. & Vázquez-Castro, M. A. Joint Network Coded-Cross Layer Optimized Video Streaming Over Relay Satellite Channel 3rd International Conference on Wireless Communications and Mobile Computing (MIC-WCMC 2013): 14-16 June 2013, Valencia, Spain, 2013.
- [12] Afanasyev, A.; Tilley, N.; Reiher, P.; Kleinrock, L., "Host-to-Host Congestion Control for TCP," *Communications Surveys & Tutorials*, IEEE , vol.12, no.3, pp.304,342, Third Quarter 2010 doi: 10.1109/SURV.2010.042710.00114.
- [13] Aumann, Robert J. and Hart, S., (1992), *Handbook of Game Theory with Economic Applications*, vol. 1, 1 ed., Elsevier.