CONVEX OPTIMIZATION BASED CONGESTION CONTROL IN LAYERED SATELLITE NETWORKS

R. R. P. Kumar[1], S. Muknahallipatna[2] and J. E. McInroy[3]

[1]National Center for Atmospheric Research, Colorado & Univ. of Wyoming, Wyoming
[2]Dept. of Electrical & Computer Engineering, University of Wyoming, Wyoming, USA
[3]Dept. of Electrical & Computer Engineering, University of Wyoming, Wyoming, USA

ABSTRACT

A multi-layered satellite network consisting of geosynchronous and nano-satellites is suited to perform space situational awareness. The nano-satellites collect information of space objects and transfer data to ground stations through the geosynchronous satellites. The dynamic topology of the network, large propagation delays and bulk data transfers results in a congested network. In this paper, we present a convex optimization based congestion control algorithm. Using snapshots of the network, operating parameters such as incoming, outgoing rates and buffer utilization are monitored. The operating parameters of a satellite are formulated as a convex function and using convex optimization techniques, the incoming data rates are evaluated to minimize congestion. Performance comparison of our algorithm with Transmission Control Protocol congestion control mechanism is presented. The simulation results show that our algorithm reduces congestion while facilitating higher transmission rates.

KEYWORDS

Congestion control, Convex Optimization, Multi-layered satellites.

1. INTRODUCTION

A Single layer satellite networks (SLSN) have the potential to provide global coverage with high bandwidth availability. SLSN can be used to provide communication infrastructure in remote areas and allow also interconnection of local area networks and individual hosts. SLSN is highly reliable due to reduced link failure instances. However, SLSN is shown to be inefficient with respect to data transmissions [1]. To address this issue, in the past two decades, multi-layered satellite network (MLSN) has been proposed by a number of researchers [2, 3].

A MLSN consisting of geosynchronous satellites and a large number of nano-satellites with different capabilities distributed across multiple layers is ideally suited to perform space situational awareness (SSA). SSA involves collecting visual information of space objects like stars, planets, satellites etc. SSA using a MLSN involves the use of nano-satellites to collect visual information and transfer the data to ground stations via multiple geosynchronous satellites through the layered network in real time. The visual information has dense data transmissions between satellites. Furthermore, the large physical distances between satellites will result in large transmission (propagation) delays, causing congestion. Packet drops due to congestion and the associated re-transmission of dropped packets makes the MLSN unsuitable for real-time SSA. Hence the need for an algorithm that can reduce the congestion by maintaining maximum possible data rates is required.

A number of satellite networks using different flavors of TCP implementations [4] have been proposed. A number of researchers [5, 6, 7, and 8] have analyzed the reduced throughput of the TCP based satellite network due to large propagation delay, slow start, packet loss assumption due to congestion. The packet loss assumption due to congestion will unduly trigger, congestion control mechanisms resulting in further throughput degradation [9]. A number of modifications to TCP congestion control mechanism (TCP-CCM) have been proposed in the past two decades [10, 11, 12, 13, 14, and 15]. The first set of modifications proposed, address either preventing congestion or reduce the congestion rapidly, while a second set of modifications (Fast Retransmit and Fast Recovery, etc.) focus on detecting whether a packet loss is due to congestion. However, the throughput of the network continues to be low during the modified congestion control mechanism operation specifically in satellite networks due to large propagation delay and bulk data transfers. In this paper, we propose a convex optimization based congestion control (COCC) algorithm that uses convex optimization to reduce congestion and achieve a maximal network throughput.

Three important parameters that contribute to congestion in a satellite network are the input buffer size, incoming and outgoing data rates of each satellite. The input buffer size and the incoming data rates determine the effectiveness of receiving data. The outgoing data rate determines the effectiveness of processing the received transmission and relaying it to the next satellite in the chain of communication links. An imbalance in the parameters can result in a satellite receiving more data than it can process, leading to congestion. The proposed algorithm consists of formulating the input buffer utilization and the incoming data rate as a convex function. Convex optimization is used to solve this convex function with associated constraints to determine the maximal possible throughput of each satellite and thereby reduce network congestion.

The paper is organized as follows: In section 2, we discuss the related work on congestion control in multi-layered satellite networks. Section 3 presents the multi-layered satellite network architecture. Section 4 provides a brief discussion of the parameters that influence congestion. In section 5, we present the traditional TCP modeling that is used for comparison. Section 6 presents the formulation of congestion control as a convex function. Section 7 presents an introduction to convex optimization and its application to congestion control. Simulation and performance evaluation of the proposed algorithm is presented in Section 8 by comparing the performance with traditional TCP-CCM. Section 9 concludes with discussion on performance issues of the COCC algorithm.

2. RELATED WORK

A QoS oriented congestion control algorithm is proposed for satellite networks in [16]. A satellite node utilizes an equation to compute the sending rate for each data flow, while the intermediate satellite nodes continuously detect real-time package-loss rates for timely adjustments. Simulation results indicate that the algorithm can provide superior congestion control performance, and raise network throughputs without reducing the QoS. However, the equation used to compute the sending rate for each data flow does not optimize the network throughput. Neither does it provide any flexible control mechanism to control source or intermediate satellite nodes to modulate sending rates. Moreover, effects of the algorithm for a multi-layered satellite network are not presented.

Congestion control using an optimized load-balancing traffic distribution algorithm for two-layered satellite network is proposed in [17]. The load-balancing scheme of the proposed method is developed by adopting a traffic distribution model, which is based upon network capacity estimation and theoretical analysis of the congestion rate in each layer. When congestion is

detected, the routing tables of satellites are modified to avoid the congested nodes. The performance of this method is effective in terms of improved throughput and lower packet drops.
Congestion control for a multi-layered satellite network in [18] is based on the probability of packet drop. Queuing ratios of satellites are varied based on the probability of packet drops at the given instance of time. This determines the traffic reduction ratio. New transmission rates are computed using the ratio that reduces congestion.

A congestion control algorithm based on sudden start and rapid recovery algorithm is introduced in [19]. The sudden start increases the transmission window size rapidly. Probe packets are transmitted periodically to check for congestion. On congestion, the rapid recovery phase algorithm cuts the window size by half for every lost packet. The performance shows higher network throughput and better fairness in sharing network resources in comparison to TCP-CCM. The limitation of the algorithm is the additional data overhead due to the probe packets.

Multilayer multicast congestion control algorithm is introduced in [20]. The satellites are grouped to retrieve session information from the ongoing traffic. The routing is computed based on the session information. Additionally, every packet is marked with priorities by every layer. Packets of lower priority are blocked during congestion and released after recovery. The algorithm has the advantage of being reliable in case of link failures, long and variable delays, limited control overhead and fair sharing of network resources.

Congestion control algorithm for lower earth orbit satellites is introduced in [21]. The round trip time (RTT) for any transmission is estimated. For a given route, the satellites are grouped based on the same number of hops and RTT. A feedback window is multicast once for every RTT to avoid congestion. The algorithm requires no modifications to a router or end-user. The performance indicates better load balancing and link utilization than traditional congestion control algorithms.

A fuzzy logic based congestion control algorithm is introduced in [22]. The algorithm formulates congestion as a function of queuing and weather characteristics. The history of weather changes and queuing for every satellite is maintained. The algorithm then computes fuzzy logic table providing the probable values of these variables. This helps in tuning of Random Early Detection (RED) algorithm. The performance of the algorithm is shown to be better than the traditional RED algorithm.

A congestion controller using data-driven switching control theory is introduced in [23, 24]. A control scheme of proportional integral-derivative structure is used to represent the congestion in networks. The controller monitors the network for any congestion. A cost function is designed to evaluate control parameters for the controller. The parameters deduced show that the algorithm is computationally less intensive than most common algorithms making it suitable for real time applications.

A study on the set of guidelines governing satellite queuing system is provided in [25]. It provides a fair routing algorithm that selectively drops packets to reduce congestion. The algorithm discriminates packets that impose bandwidth more than their allocation. This discrimination enables the satellite to drop the right packets during congestion. The performance is shown to be better than traditional congestion control algorithms through simulations. However, Huang *et al* state in [25] that the implementation on a satellite network may not be feasible.

3. MOTIVATION

As mentioned before, SSA using MLSN involves a large number of nano-satellites, with each satellite involving dense data transmissions. Furthermore, a real-time SSA using MLSN would require to have maximum feasible network throughput, even during congestion phase.

In all of the previous work discussed, the congestion is reduced by reducing the transmission rate either linearly or exponentially without any consideration to the network throughput. The review of congestion control algorithms in [26] shows the same. Our proposed algorithm differs significantly by adopting a different goal for congestion control. The goal is to clear congestion while maintaining maximal network throughput. To achieve this goal, congestion control is formulated as a convex function with incoming data rates of each satellite as the variable of optimization. It reduces congestion and optimizes the data flow simultaneously, providing an efficient network throughput.

4. ARCHITECTURE

A novel multi-layered satellite routing algorithm is proposed in [27]. The proposed routing algorithm performance is demonstrated on a satellite network consisting of satellites distributed over multiple layers with an individual layer situated either at lower earth orbit (LEO) or middle earth orbit (MEO) or geosynchronous earth orbit (GEO). The performance of the network in [27] is shown to have low communication overhead and better throughput than other fewer-layered satellite networks. However, the focus of this work is only on optimal routing and does not address the issues with congestion and maximal network throughput. In this work, we demonstrate our convex optimization based congestion control algorithm on the satellite network test bed, which is a modified version of the layered satellite network architecture proposed in [27]. The modifications are the LEO and the MEO layers, are referred to as layer-1 and layer-2 respectively, comprising of only nano-satellites. Layer-1 and layer-2 are not expected to be situated at the low and medium earth orbits. The GEO layer is referred as layer-3 consisting of GEO satellites capable of communicating with the ground stations. The communication (transmission and reception rates) capabilities of satellites are assumed to increase from layer-1 through layer-3. Satellites are assumed to communicate within and between layers via intra-orbital and inter-orbital links respectively. It is assumed that every satellite knows its position via geographic coordinate system. The hierarchy of the satellites is shown in Fig 1.

Let the number of GEO satellites be N_G, number of layer-2 satellites be N_M and number of layer-1 satellites be N_L. The satellites are represented by

$$G = \{g_i \mid i = 1,2,...,N_G\} \tag{1}$$

$$M = \{m_j \mid j = 1,2,...,N_M\} \tag{2}$$

$$L = \{l_k \mid k = 1,2,...,N_L\} \tag{3}$$

where, g_i, m_j, l_k represent the individual layer-3, layer-2 and layer-1 satellites respectively. As seen in Fig. 1, layer-1 has two sub-layers deviating from the architecture proposed in [27]. In order to efficiently maintain data flow between satellites, manager or cluster head (CH) satellites are introduced.

The individual satellite naming convention used in identifying the data flow or links is discussed below:

- For layer-1, satellite links have two representations, $l_{k,h}$ and $l^h_{k,j}$, where $l_{k,h}$ is the link between k^{th} non-cluster head satellite in layer-1 and CH satellite h of layer-1. $l^h_{k,j}$ is the satellite link between k^{th} CH satellite in layer-1 to the j^{th} satellite in layer-2.

- Satellites ($m_{j,i}$) in layer-2 are arranged in a single orbit, where j is the identifier of a satellite in layer-2 and i is the identifier of a satellite in layer-3.
- Layer-3 will have satellites that may or may not have ground connectivity. The satellites having connectivity to a ground station are selected as cluster heads. The two identifiers of satellite links in this layer are $g_{i,h}$ and g^h_i. $g_{i,h}$ is the link between i^{th} satellite and CH satellite of layer-3. g^h_i is the link between the i^{th} CH in layer-3 to the ground station.

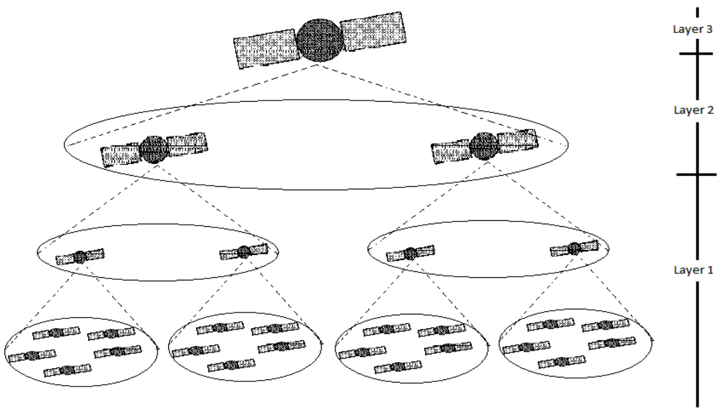

Fig. 1. Hierarchy of the Multi-layered Satellite Network

The data collected by $l_{k,h}$ is transmitted to $l^h_{k,j}$. Each $l^h_{k,j}$ routes this data to $m_{j,i}$, which further routes it to either $g_{i,h}$ or g^h_i. The $g_{i,h}$ relays their data to g^h_i, and eventually data to the ground stations.

5. CONGESTION CONTROL PARAMETERS

Due to the dynamic topology of a satellite network, the parameters like propagation delay, maximum possible data rates, etc., change within a well-defined bandwidth. However, these parameters can be assumed to be constant within a snapshot. A snapshot is defined as a brief period of time and in which the network topology change is minimal. A snapshot approach is useful in analyzing the current state of a dynamic network, and determines the operational parameters of the network for the next state. At the beginning of every snapshot, every satellite g_i, m_j and l_k based on their current position, will compute the following three parameters:

- Line of sight with other satellites,
- Maximum data transmission rates
- Data recipients

5.1. Line of Sight

The inter-orbital and intra-orbital links being wireless require a LOS for transmission. The satellites using the geographical coordinate system will determine the LOS satellites as discussed in [28, 29]. To determine LOS, position vectors of a satellite and difference vectors are used. A position vector of a satellite is the Euclidean vector representing the position of the satellite with the center of the earth as its origin. A difference vector is the Euclidean vector obtained by the subtraction of two Euclidean vectors. Using these vectors, LOS is computed as follows:

- Let θ_1 represents the angle between satellite A's position vector and the difference vector (difference between A and B satellites' position vector).
- Let θ_2 represents the angle between satellite B's position vector and the difference vector.
- A LOS exists between A and B if any of the following conditions is satisfied:

  - ❖ $\theta_1 > 90^\circ$
  - ❖ $\theta_1 \le 90^\circ$ and $\theta_2 < 90^\circ$
  - ❖ $\theta_1 \le 90^\circ$, $\theta_2 \ge 90^\circ$ and the orthogonal from the center of the earth to the line joining the two satellites is greater than the radius of the earth.

5.2. Maximum Data Transmission Rates

Once the LOS between two satellites is determined, the satellites are considered as neighbors. Laser transmission is assumed as the mode of communication in this work to achieve high transmission rates. Laser transmission rate is dependent on a number of parameters [30] like transmission power, area of transmitter antenna, distance between satellites, etc. For a given laser communication configuration, the relationship between maximum data transmission rate and the distance between two satellites can be expressed as

$$R_{\max} \alpha \frac{1}{D^2} \tag{4}$$

where,

D is the distance between the two satellites.

D varies between satellites in different layers constantly due to their orbital locations. The transmission rate of a transmitting satellite is the arrival rate at the receiving satellite. Even though, $R_{\max}$ is the maximum possible transmission rate of a satellite in a snapshot, the actual transmission rate will be dictated by the underlying TCP.

5.3. Data Recipients

A top down approach is adopted to select the data recipients at each layer. The notations and data flow for the network is as shown in Fig. 2. CH satellites are primarily chosen based on LOS. Satellites that do not have a LOS with a CH in a snapshot, do not participate in any transmission activity. Satellites having LOS to multiple CHs can choose to transmit to any or all of them.

CH g_i^h is chosen based on its connectivity to a ground station. CH $l_{k,j}^h$ is chosen based on greedy algorithm of maximum neighboring l_k satellites. From a set of l_k satellites having a $m_{j,i}$ neighbor, a satellite in the set with maximum number of neighboring l_k is chosen as a CH.

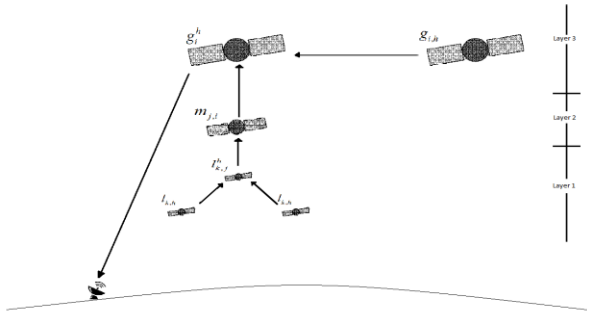

Fig. 2. Data flow in the Multi-Layered Satellite Network

6. TCP MODELLING

In this work, the TCP flavor described in [31] has been adopted for comparison with the convex optimization based congestion control algorithm.

At the beginning of every snapshot, the satellites resume transmission activity based on the computed hierarchy of transmission. The allocated rate of transmission R_{alloc}, governed by TCP, will be a fraction of $R_{\max}$ as given in Eq. 5

$$R_{alloc} = T_{tcp} R_{\max} \tag{5}$$

where, T_{tcp} is the a threshold satisfying Eq. 6

$$0 \le T_{tcp} \le 1 \tag{6}$$

The threshold T_{tcp} varies based on the success or failure of every transmission. In every snapshot, the satellites follow the TCP principles i.e., slow start or exponential growth phase below a preset TCP transmission rate and a linear growth thereafter. The linear growth is continued till congestion is detected or T_{tcp} reaches unity. If congestion is detected, the transmission is subjected to the TCP-CCM where the slow start window is cut by half and slow start is restarted. Once the congestion is cleared, the TCP resumes back its transmission with linear growth. If no congestion is detected, the transmission rate is maintained at $R_{\max}$. Furthermore, the satellites are allowed to transmit data in bursts. The amount of data a satellite can transmit in bursts is limited by the bandwidth-delay product [32].

7. FORMULATION OF CONGESTION CONTROL AS CONVEX FUNCTION

Let the service rate of packets of a satellite be μ. Let the maximum arrival rate of packets on an i^{th} input link of a satellite be λ_i. In order to empty the buffer, and thus reduce congestion, the service rate and the arrival rates must satisfy the relation:

$$\mu > \sum_{i=1}^{n} \lambda_i \tag{7}$$

where, n indicates the number of satellites transmitting. At $\mu = \sum_{i=1}^{n} \lambda_i$, the network is said to be at a critical state indicating that the buffer is always empty. For any satellite with buffer capacity B, Eq. 7 can be re-written as:

$$\mu > \sum_{i=1}^{n} \lambda_i + (B_{util} \times B) \tag{8}$$

where,

$0 \le B_{util} \le 1$.

The product $B_{util}B$ represents buffer utilization. If the data rates in Eq. 8 are optimized, a network will operate with minimal congestion and maximal throughput.

Let $R_{\max}^{n} \in R^n$ be maximum transmission rates of all satellites transmitting to a single CH satellite:

$$R_{\max}^{n} = \begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_n \end{bmatrix} \tag{9}$$

where, n is the number of incoming links of a CH satellite. In order to find an effective data rate satisfying the condition of minimum queuing delay, a scalar multiple for each λ_i has to be considered. The collection of scalar multiples is represented as $W = \begin{bmatrix} w_1 & w_2 & \dots & w_n \end{bmatrix}$, where $0 \le w_i \le 1$, such that:

$$(R_{\max}^{n})^T W < \mu - (B_{util} \times B) \tag{10}$$

Eq. 10 is a linear equation of $W \in R^n$. Next, we will show Eq. 10 to be a convex function.

8. Convex Optimization And Application To Congestion Control

A set S, is defined as convex if and only if it satisfies the condition [33] described in Eq. 11.

$$\theta x_1 + (1-\theta) x_2 \in S \tag{11}$$

where,

$x_1, x_2 \in S$,

$x_1 \ne x_2$,

$\theta \in R$, and

$0 \le \theta \le 1$.

A function $f(X): R^n \to R$ is considered convex [33] if and only if for all $X_1, X_2 \in domain(f)$ satisfies Eq. 12.

$$f(\theta X_1 + (1-\theta) X_2) \le \theta f(X_1) + (1-\theta) f(X_2) \tag{12}$$

Consider the inequality,

$$a^T X \le b \tag{13}$$

where,

$a \in R^n$,

$b \in R$, and

$X \in R^n$ is the unknown to be determined.

Eq. 13 represents a linear inequality in n-dimensional half-space with multiple solutions. X is a convex set since it satisfies Eq. 11 and therefore is a solution of Eq. 13 for a given a, b. S. Boyd and L. Vanderberghe have shown linear functions to be convex functions [33].Comparing Eq. 10 and Eq. 13, it can be seen that $\left(R_{\max}^{n}\right)^{T}$ and $\mu-\left(B_{util}\times B\right)$ are a^{T} and b respectively, thereby Eq. 10 is a convex function and a linear inequality.

Eq. 10 or Eq. 13 can also be solved using a simplex algorithm. Simplex algorithms are infeasible for large data [34]. For SSA using MLSR, a large number of nano-satellites are involved. Hence, in Eq. 10, as n grows large, simplex algorithm becomes infeasible. Furthermore, simplex algorithm optimizes linear functions only. Convex optimization can be applied on a convex function which can be linear, quadratic or geometric. Hence, optimizing congestion control formulated as a convex function allows future work to add additional parameter or constraints to congestion control for different or complex networks.

Any standard convex optimization toolkit can be used to solve Eq. 10. To obtain the global minima and in this work, the CVX toolkit [33] developed by S. Boyd and L. Vanderberghe for implementing convex optimization in Matlab is used. To solve a convex function, this toolkit requires the convex function to be specified in a particular format. The toolkit requires objective variable and convex constraints to solve the optimization problem. The objective in the congestion control problem is optimizing W which is n-dimensional. However, the toolkit allows the objective to be only a single dimension variable. Therefore, a variable γ is introduced which will be maximized for each satellite as shown in Eq. 14.

$$
\begin{aligned}
&\text{maximize}\,\gamma \\
&\text{subject to} \\
&\qquad \left(R_{\max}^{n}\right)^{T} W < \mu-\left(B_{util}\times B\right) \\
&\qquad 0 \le w_i \le 1 \\
&\qquad w_i \lambda_i \ge \alpha_i \\
&\qquad w_i \lambda_i \le \beta_i \\
&\qquad \gamma < w_i
\end{aligned}
\tag{14}
$$

where,

α_i, β_i are the lower and upper limits of the ith incoming link rate.

Since γ faces the constraint $\gamma < w_i$, in-turn, all elements of w_i are maximized. The constraint $0 \le w_i \le 1$ forces w_i to be a convex set. The other two constraints are to enforce the effective incoming link rate is maintained within a bandwidth. The advantage of convex optimization is observed in these two constraints. It facilitates to optimize transmission rates within the desired bandwidth which would not be possible if least squares technique was used.

The COCC algorithm on a CH satellite is triggered when the effective buffer utilization is within the bounds defined by $[B_{lower}, B_{upper}]$. Therefore, the COCC algorithm is activated at

$B_{util}B > B_{upper}$ and deactivated at $B_{util}B < B_{lower}$. Once the COCC phase ends, TCP-CCM regains control of the transmission.

A CH after computing the optimal transmission rates $(\lambda_i w_i)$ for its incoming links, relays these desired rates to all of its neighboring satellites and this process is repeated on all congested CH satellites.

9. Simulations And Analysis

The simulations are tailored to evaluate performance of COCC in comparison to TCP. The comparison is performed by considering two parameters namely:

- Average Buffer utilization of a Layer: Average Buffer Utilization is defined as the average of the buffer utilization of each CH satellite in a layer for a snapshot.
- Average Link utilization of a Layer: Average Link Utilization is defined as the average of the link utilization of each satellite in a layer for a snapshot.

Simulations are performed with 135 satellites distributed across the three layers in the ratio of 1:4:40. Layer-1 contains 120 satellites distributed in orbits with altitude ranging from 28,000 km to 32,500 km. Layer-2 contains 12 satellites distributed in orbits with altitude ranging from 33,000 km to 35,000 km. Layer-3 contains 3 satellites in the geosynchronous orbit. Based on the above satellite distribution, the average theoretical data rates for a CH satellite in layer-1, layer-2 and layer-3 was 1 Mbps, 10 Mbps and 8 Mbps respectively. The layer-1 theoretical data rate is significantly less compared to layer-2 theoretical data rate due to a smaller antenna with lower transmission capability. Layer-3 theoretical data rate is also lower compared to layer-2 theoretical data rate due to the large distance between geosynchronous satellites and ground stations and the effect of earth's atmosphere on laser transmission. The buffer size of each satellite in layer-1, layer-2 and layer-3 were set to 1 MB, 1 GB and 10MB respectively. The buffer size on a geosynchronous satellite was set to a lower value to increase the effect of congestion. Initially, simulations were performed with a 30 second snapshot intervals. Since our proposed algorithm works only on network layer, the network flow for duration of 30 minutes was simulated using the CVX toolkit on Matlab and Satellite ToolKit (STK). The simulation results presented are based on 10 trials. The standard deviation from the 10 trials for average buffer utilization and average link utilization vary from 3.7% to 6.7% across the layers.

In Figs. 3, 4 and 5 the average buffer utilization for layers 3, 2 and 1 are shown respectively. In Figs. 6, 7 and 9 the corresponding average link utilization for layers 3, 2 and 1 are shown respectively.

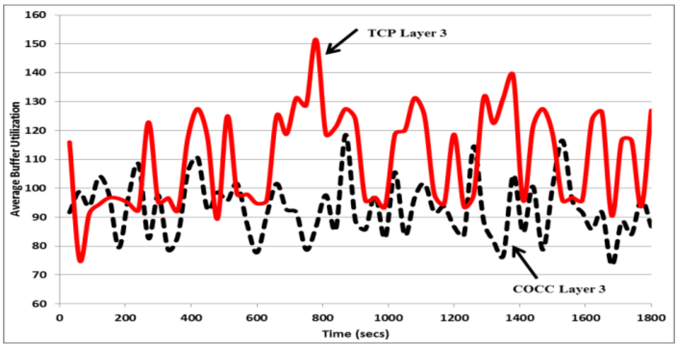

Fig. 3. Layer-3 Average Buffer Utilization for TCP-CCM and COCC

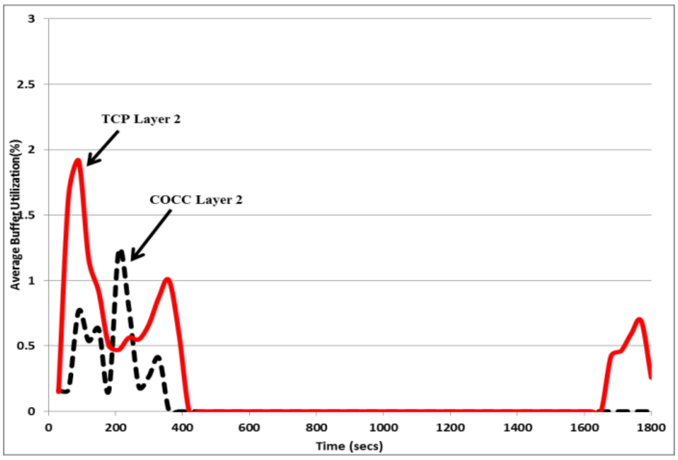

Fig. 4. Layer-2 Average Buffer Utilization for TCP-CCM and COCC

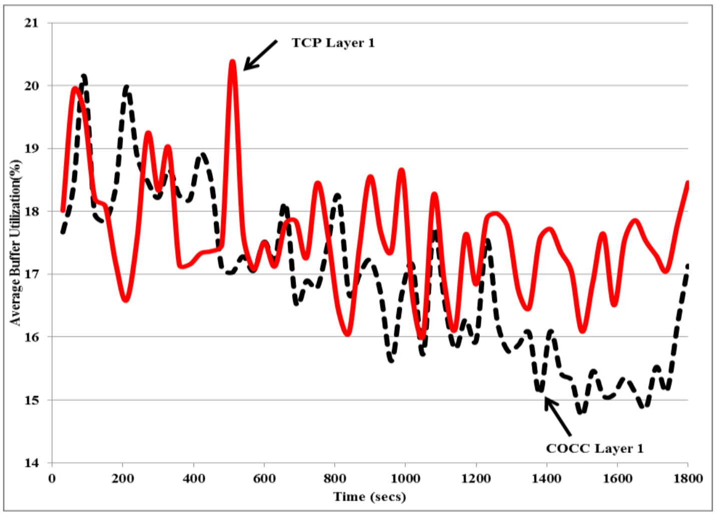

Fig. 5. Layer-1 Average Buffer Utilization for TCP-CCM and COCC

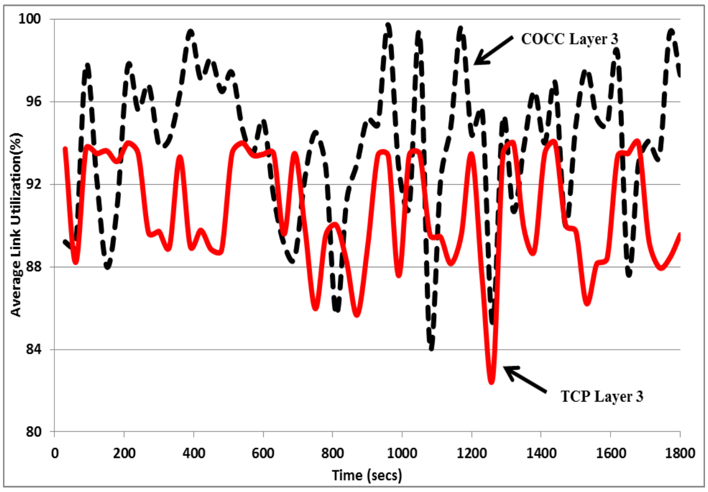

Fig. 6. Layer-3 Average Link Utilization for TCP-CCM and COCC

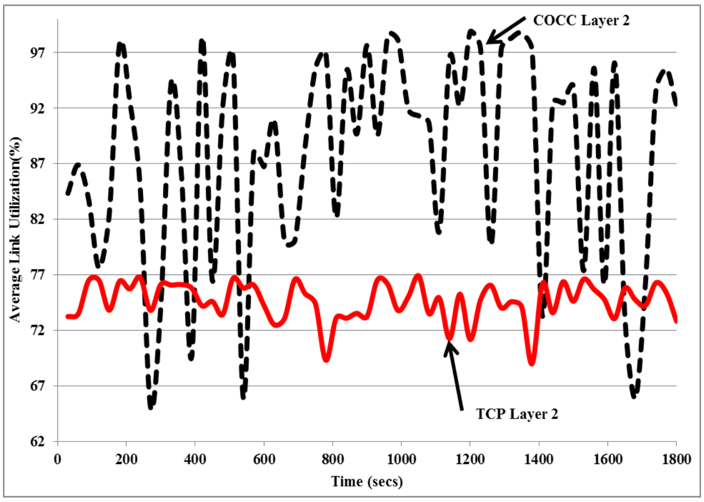

Fig. 7. Layer-2 Average Link Utilization for TCP-CCM and COCC

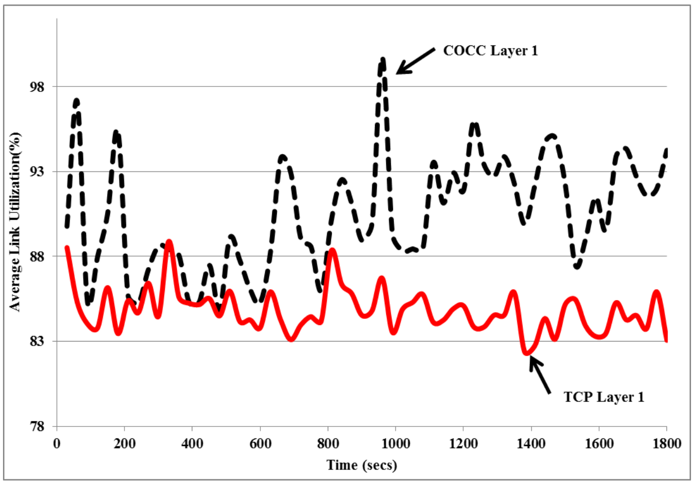

Fig. 8. Layer-1 Average Link Utilization for TCP-CCM and COCC

In Fig. 3 the average buffer utilization of satellites in layer-3 with traditional TCP-CCM is above 100% on an average for the entire simulation. This indicates significant number of packets being dropped. However, with the COCC algorithm on an average, the buffer utilization is around 100%. Since the buffer utilization at layer-3 with TCP-CMM is above 100%, the corresponding average link utilization is less than 80% for satellites in layer-2 as seen in Fig. 7. However, with COCC algorithm, the average link utilization for the same satellites in layer-2 is close to 90% due to buffer utilization reduction in layer-3. In Fig. 4, the buffer utilization with TCP-CCM or COCC is negligible compared to other layers. This is due to the satellites in this layer having higher theoretical transmission rate, larger buffer size and experiencing a lower input data rate from CHs in layer-1. It was observed due to the layer-2 and layer-1 orbit altitudes, in any given snapshot, a total of 30 satellites among the 120 satellites in layer-1 were chosen as CHs. Hence each CH in layer-1 is servicing a maximum of 3 non-CH satellites. Therefore the average buffer utilization of the CHs in layer-1 is only around 20% for both TCP-CCM and COCC as seen in Fig. 5. Since the ground stations do not experience any congestion in our simulations, the link utilization in layer-3 is high as seen in Fig. 6. The link utilization of layer-1 satellites is not affected by the buffer utilization in layer-2. The link utilization with TCP-CCM in layer-1 is less compared to that with COCC as seen in Fig. 8. This is due to the use of slow start and congestion avoidance by TCP-CCM at the start of a new snapshot.

Even though COCC performs better congestion control it still has some limitations. One of the limitation is the lethargic congestion control at activation of COCC. When COCC is activated on a congested CH satellite, the new transmission rates to reduce congestion is computed and transmitted on all its incoming links with a delay. Due to this delay, the buffer utilization is still shooting above 100% in Fig. 3. This limitation can be overcome by allowing COCC a continuous control over congestion control mechanism. However, allowing COCC to perform continuous congestion control imposes a large computational burden as shown in Fig. 11. To avoid this computational burden, instead of allowing COCC to perform continuous congestion control we have explored varying the duration of a snapshot. In Fig. 9, the buffer utilization with snapshot durations of 30 and 15 seconds are shown. It can be seen the buffer utilization not exceeding 100% with 15 seconds snapshot and thereby no packets are dropped resulting in high link utilization as seen in Fig. 10.

As previously mentioned, the other limitation is the computational burden of COCC. The computation time of COCC is dependent on parameters such as number of neighboring satellites, snapshot interval, data transmission rates, iterations involved in convex optimization etc. For a 30s snapshot interval executed on a Quad-core desktop, an average computation time was 3.25 minutes with 50 satellites as shown in Fig. 11. It can also be noticed that the computational time increases linearly with increasing number of satellites. At this stage due to the computational burden COCC is not suitable for real-time application.

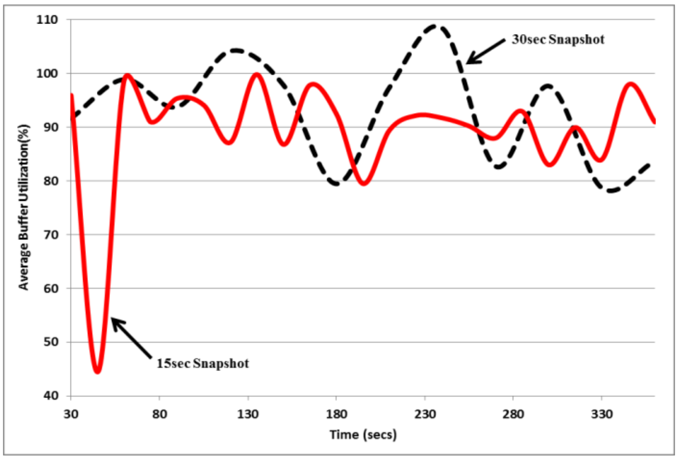

Fig. 9. Layer-3 Average Buffer Utilization for varied Snapshot Interval

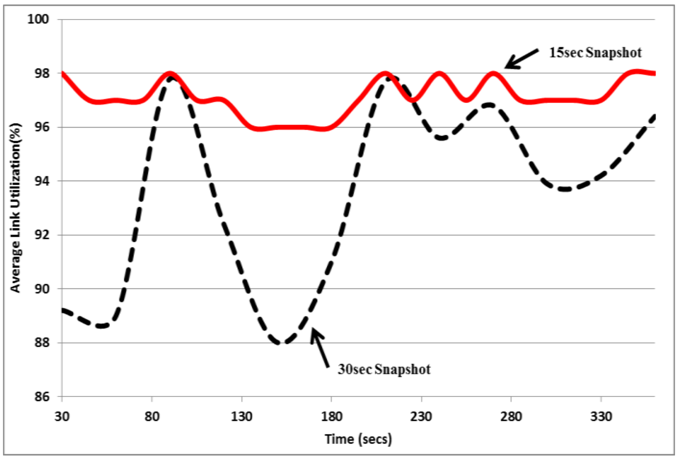

Fig. 10. Layer-3 Average Link Utilization for varied Snapshot Interval

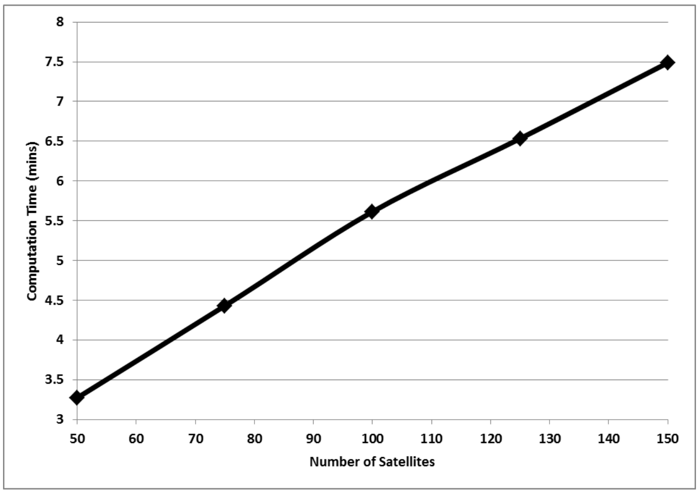

Fig. 11. New Data Rates Computation Time with COCC

10. CONCLUSIONS

The paper proposes a new congestion control algorithm for layered satellite networks. The approach is to formulate congestion control as a convex optimization problem. The convex function is optimized using convex optimization approach at discrete intervals of time to determine optimal transmission rate of each satellite such that the network congestion is reduced while maintaining optimal network throughput. The performance of COCC was compared with TCP-CCM and the performance of COCC is better. It was observed that the satellites transmitted higher transmission rates with COCC algorithm. Furthermore, the performance of COCC is improved when the snapshot duration is reduced. Currently, a constant control by COCC is not feasible due to large computational burden.

To use COCC algorithm in real-time, the COCC algorithm needs to be executed in parallel for each satellite necessitating parallelization of the CVX toolkit.

ACKNOWLEDGEMENTS

We would like to thank National Center for Atmospheric Research, Boulder, for providing resources.

REFERENCES

[1] Werner, Markus, Cecilia Delucchi, Hans-Jörg Vögel, Gerard Maral, & Jean-Jacques De Ridder, (1997) "ATM-based routing in LEO/MEO satellite networks with intersatellite links", IEEE Journal on Selected Areas in Communications, Vol. 15, No. 1, pp. 69-82.

[2] Lee, Jaeook, & Sun Kang, (2000) "Satellite over Satellite (SOS) Network: A Novel Architecture for Satellite Network", Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, Vol. 1, pp. 392–399.

[3] Nishiyama, Hiroki, Daigo Kudoh, Nei Kato, & Naoto Kadowaki, (2011) "Load Balancing and QoS Provisioning Based on Congestion Prediction for GEO/LEO Hybrid Satellite Networks", Proceedings of the IEEE, Vol. 99, No. 11, pp. 1998–2007.

[4] Wang, Ruhai, Tarik Taleb, Abbas Jamalipour, & Bo Sun, (2009) "Protocols for reliable data transport in space internet", IEEE Communications Surveys Tutorials, Vol. 11, No. 2, pp. 21-32.

[5] Lakshman, T. V., & Upamanyu Madhow, (1997) "The performance of TCP/IP for networks with high bandwidth-delay products and random loss", IEEE/ACM Transactions on Networking, Vol. 5, No. 3, pp. 336-350.

[6] Metz, Christopher, (1999) "TCP over satellite… The final frontier", IEEE Internet Computing, Vol. 3, No. 1, pp. 76–80.

[7] Partridge, Craig, & Shepard, Timothy, (1997) "TCP/IP performance over satellite links", IEEE Network Mag., pp. 44–49.

[8] Bakre, Ajay, and B. R. Badrinath, (1995) "I-TCP: Indirect TCP for mobile hosts", International Conference on Distributed Computing Systems (ICDCS), pp. 136–143.

[9] Oueslati-Boulahia, Sara, Ahmed Serhrouchni, Samir Tohmé, S. Baier, & Marouane Berrada, (2000) "TCP over satellite links: Problems and solutions", Telecommunication Systems, Vol. 13, pp. 199-212.

[10] Keshav, Srinivasan, & Samuel P. Morgan, (1997) "SMART retransmission: performance with overload and random losses", IEEE Infocom, Vol. 3, pp.1131–1138.

[11] Balakrishnan, Hari, Venkata N. Padmanabhan, Srinivasan Seshan, & Randy H. Katz, (1997) "A comparison of mechanisms for improving TCP performance over wireless links", IEEE/ACM Transactions on Networking, Vol. 5, No. 6, pp. 756-769.

[12] Padmanabhan, Venkata N., & Randy H. Katz, (1998) "TCP Fast Start: A technique for speeding up web transfer", IEEE Globecom.

[13] Akyildiz, Ian F., Giacomo Morabito, & Sergio Palazzo, (2001) "TCP-Peach: A New Congestion Control Scheme for Satellite IP Networks", IEEE/ACM Transactions on Networking, Vol. 9, No. 3, pp. 307-321.

[14] Utsumi, Satoshi, Salahuddin Muhammad Salim Zabir, & Shiratori, Norio (2008) "TCP-Cherry: A new approach for TCP congestion control over satellite IP networks", Computer Communications, Vol. 31, No. 10, pp. 2541-2561.

[15] Kim, Jaehyun, Choo, Hyunseung & Kim, Dongsoo (2012) "TCP CAE: an improved congestion control using comparative ACK-based estimator", The Journal of Supercomputing, Vol. 59, No. 2, pp. 1019-1034.

[16] Liu, Heyu, & Sun, Fuchun, (2014) "A QoS-Oriented Congestion Control Mechanism for Satellite Networks", Mathematical Problems in Engineering.

[17] Nishiyama, Hiroki, Tada, Yasunori, Nei Kato, Yoshimura, Naoko, Morio Toyoshima, & Kadowaki, Naoto, (2013) "Toward optimized traffic distribution for efficient network capacity utilization in two-layered satellite networks", Vehicular Technology IEEE Transactions, Issue 62, No. 3, pp. 1303-1313.

[18] Taleb, Tarik, Mashimo, Daisuke, Jamalipour, Abbas, Hashimoto, Kazuo, Yoshiaki Nemoto, & Nei Kato, (2006) "SAT04-3: ELB: An Explicit Load Balancing Routing Protocol for Multi-Hop NGEO Satellite Constellations", Global Telecommunications Conference, pp. 1-5.

[19] Akyildiz, Ian F., Morabito, Giacomo, & Palazzo, Sergio, (2001) "TCP-Peach: a new congestion control scheme for satellite IP networks", IEEE/ACM Transactions on Networking, Vol. 9, No. 3, pp. 307-321.

[20] Peng, Jun, & Sikdar, Biplab, (2004) "Multilayer multicast congestion control in satellite environments", IEEE Journal on Selected Areas in Communications, Vol. 22, No. 3, pp. 449-461.

[21] Taleb, Tarik, Kato, Nei & Nemoto, Yoshiaki, (2006) "REFWA: An Efficient and Fair Congestion Control Scheme for LEO Satellite Networks", IEEE/ACM Transactions on Networking, Vol. 14, No. 5, pp. 1031-1044.

[22] Vallamsundar, Banupriya, Zhu, Jiesheng, K. Ponnambalam, Huang, Changcheng, Srinivasan, Anand & Cheng, Brian , (2007) "Congestion Control for Adaptive Satellite Communication Systems Using Intelligent Systems", International Symposium on Signals, Systems and Electronics, pp. 415-418.

[23] Cao, Jinhua, & Stefanovic, Margareta, (2009) "Switching congestion control for satellite TCP/AQM networks," American Control Conference, pp. 4842–4847, June 2009.

[24] Cao, Jinhua, & Stefanovic, Margareta, (2010) "New results in stable switching congestion control for satellite TCP/AQM networks", 49th IEEE Conference on Decision and Control (CDC), pp.5643-5648.

[25] Huang, Xiao, & Jiao, Chengbo, (2008) "Congestion Control in Satellite Networks", 4th International Conference on Wireless Communications, Networking and Mobile Computing, pp.1-4.

[26] Kushwaha, Vandana, (2013) "A review of router based congestion control algorithms", International Journal of Computer Network and Information Security (IJCNIS), Vol. 6, No. 1,pp. 1.

[27] Akyildiz, Ian F., Ekici, Eylem & Bender, Michael D., (2002) "MLSR: a novel routing algorithm for multilayered satellite IP networks", IEEE/ACM Transactions on Networking, Vol. 10, No. 3, pp. 411-424.

[28] Gerdan, George P., & Deakin, Rodney E., (1999) "Transforming Cartesian Coordinates X, Y, Z to Geographical Coordinates ", Australian Surveyor, Vol. 44, No. 1, pp. 55-63.

[29] Shamos, Michael Ian, & Hoey, Dan, (1975) "Closest-point problems", 16th Annual Symposium on Foundations of Computer Science, pp. 151-162.

[30] Majumdar, Arun K, (2005) "Free-space laser communication performance in the atmospheric channel", Journal of Optical and Fiber Communications Reports, Vol. 2, No. 4, pp. 345-396.

[31] Ghani, Nasir, & Dixit, Sudhir, (1999) "TCP/IP enhancements for satellite networks", IEEE Communications Magazine, Vol. 37, No. 7, pp. 64-72.

[32] Chen, Kai, Xue, Yuan, Shah, Samarth H., & Nahrstedt, Klara, (2004) "Understanding Bandwidth-delay Product in Mobile Ad Hoc Networks", Computer Communications, Vol. 27, No. 10, pp. 923-934.

[33] Boyd, Stephen, & Vandenberghe, Lieven , (2004) "Convex Optimization", Cambridge University Press.

[34] Beasley, John E., & Beasley, John E. eds, (1996) "Advances in linear and integer programming", Oxford: Clarendon Press.