Congestion Management Techniques for Future Satellite Networks

Pablo G. Madoery, Carleton University, Canada
Gunes Karabulut Kurt, Carleton University, Canada
Halim Yanikomeroglu, Carleton University, Canada
Peng Hu, National Research Council, Canada
Khaled Ahmed, Satellite Systems, MDA, Canada
Guillaume Lamontagne, Satellite Systems, MDA, Canada

Abstract

Future satellite networks will make intensive use of inter-satellite links and computing edge capabilities to conform communication networks in space that connect millions of users on the ground. Unlike terrestrial Internet-based networks, these new satellite networks will need to incorporate novel techniques to avoid or mitigate congestion. In this paper we describe the main features of schemes incorporating these techniques and compare their performance by means of simulations of a realistic satellite constellation. The results show the benefits of moving from reactive feedback-based schemes to proactive schemes that prevent congestion before it occurs.

1. Introduction

Internet access is a basic need for social and economic development. However, such access is currently inequitable because remote and sparsely populated regions do not represent a profitable business for telecommunication companies, which face challenges in establishing and maintaining an adequate communications infrastructure [1].

In order to overcome these challenges, the new space-age companies, like Starlink and Telesat, are planning to launch satellite mega-constellations to provide global broadband low-latency Internet to the far corners of the Earth. In order to unlock their true potential, satellites have begun to incorporate inter-satellite links and in-orbit processing capability. This allows the conformation of communication networks in space capable of forwarding terrestrial traffic in a similar way as when using the Internet [2].

However, one of the fundamental differences between these networks lies in the stability of the links. While Internet links are stable and with low losses, in satellite networks the links suffer frequent interruptions caused by orbital dynamics, power limitations and resource constraints. This has a drastic consequence from a networking point of view: routing and transport protocols designed for the Internet perform poorly if no adaptation is made to operate in a highly dynamic and time-varying environment as it is the space [3]. One of the problems affecting performance, which we will see here, is congestion.

2. Local Congestion vs Global Congestion

In general, the congestion problem has been defined as the attempt to send more data than what a given link or node buffer allows for. Hence, congestion is caused by a combination of topology constraints and excessive network traffic. We distinguish here between 2 types of two or more traffic sources competing for the same resources. Fig. 1 shows these two cases of congestion in a network of 3 satellites (N1, N2, N3) that have links whose availability varies over time. These variations are captured by means of states (s1, s2, s3) that represent "snapshots" of the topology at different times.

In the case of local congestion, N1 generates 10 traffic packets destined to N3 and which are sent using the route formed by the links N1-N2 and N2-N3. However, since the N2-N3 link has half the capacity of N1-N2, only 5 packets will be able to pass through to the destination. Here, the congestion problem is caused because N1 only considered the capacity of the first link, and not that of the successive links on the route.

In the case of global congestion, all links have the capacity to send 10 traffic packets. However, this time both N1 and N2 generate 10 traffic packets destined for N3. Therefore, when the 10 packets from N1 reach N2, it will only be possible to deliver half of the generated traffic to the destination. Here the congestion problem is caused because N1 is not aware of the traffic generated by intermediate nodes on a route.

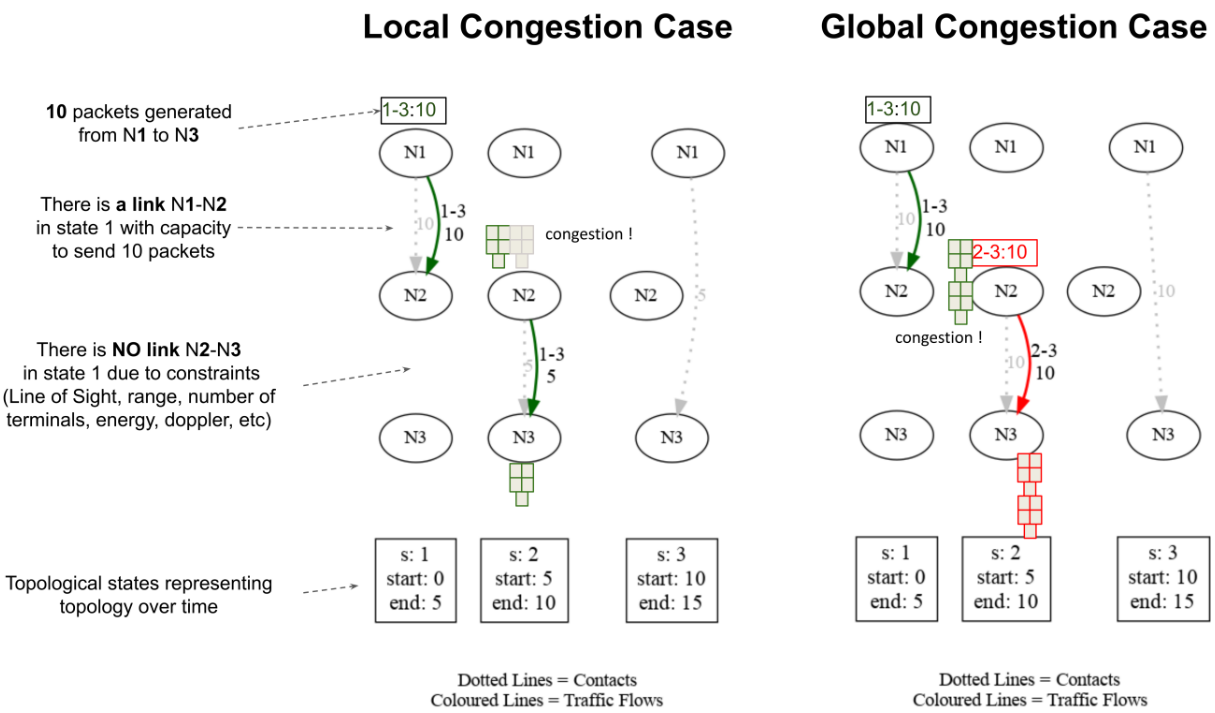

Fig. 1: Comparison between a case of local congestion, due to a single traffic source, and a case of global congestion, due to 2 traffic sources.

3. Reactive solutions vs Proactive solutions

In traditional Internet-based terrestrial networks, protocols such as Transmission Control Protocol (TCP), are extensively used to avoid or mitigate congestion. This protocol works on the assumption that there is a stable path, perhaps including multiple links/hops, from the source to the destination. The destination uses that stable path to send feedback (acknowledgments) to the source as it receives the sent traffic. When the source does not receive the appropriate acknowledgments, it infers that it is due to congestion, and reduces the transmission data rate to mitigate that congestion. This *reactive* operation does not work efficiently when the lack of appropriate acknowledgements is due to handovers or link switching, rather than actual congestion. Because the source is not able to distinguish between the two cases, it ends up reducing the data rate unnecessarily, affecting network performance [3]. Alternatively, several proactive schemes have also been proposed [4][5] that leverage knowledge of topology and expected future traffic as a means of avoiding congestion before it occurs. In this paper we will compare the following schemes that incorporate different congestion avoidance capabilities in a well-known routing algorithm called CGR [6]:

CGR (Contact Graph Routing): Each satellite receives *the same* "contact plan" containing the future topology. When traffic is generated or received, the routing algorithm is executed at each satellite in a distributed manner considering only the capacity of the link to the next neighbor on a route.

LPA-CGR_FB (Local Path Aware CGR): It works similarly to CGR, but when routing traffic, each satellite considers the capacity of all the links in a route (not just the first one).

GPA-CGR (Global Path Aware CGR): Each satellite receives a *different* contact plan, which restricts the topology with an allocation algorithm in such a way as to avoid congestion when future traffic is known.

EGPA-CGR (Evolutionary Global Path Aware CGR): It works similarly to GPA-CGR but the allocation algorithm used to constrain the topology is a genetic algorithm.

4. Evaluation and Results

To evaluate the described congestion avoidance schemes, we study a realistic linear constellation of 4 low Earth orbit satellites like the one proposed by NASA's A-Train constellation [7]. The constellation is propagated over a 24-hour time interval to obtain the time-varying topology, which is then fed to a discrete-event simulator called Dtnsim [8] that was extended to incorporate the schemes described here.

The main metrics calculated are the *delivery ratio*, which expresses the percentage of packets that reach their destination, and the *link usage per packet delivered*, which shows how efficiently the network's links are used. Each of these metrics is calculated for a variable traffic load ranging from 0 to 1 on the abscissa axes.

When the traffic load is low, all schemes show the same delivery ratio. However, when the traffic load increases beyond 0.3, CGR begins to experience a noticeable drop in delivery ratio due to congestion. LPA-CGR has intermediate performance due to its consideration of the capacity of intermediate links, while schemes that consider globally generated traffic have the best performance due to their ability to proactively prevent congestion.

With respect to how efficient these schemes are in the use of resources, CGR is the least efficient since it has the highest link utilization per packet delivered. Packets get stuck at intermediate nodes and end up making multiple hops until they finally reach their destination. On the other hand, LPA-CGR avoids these bottlenecks to some extent and is further improved by both GPA-CGR and EGPA-CGR. These latter schemes leverage the knowledge of a global view of the network in terms of topology and traffic to make better use of the links.

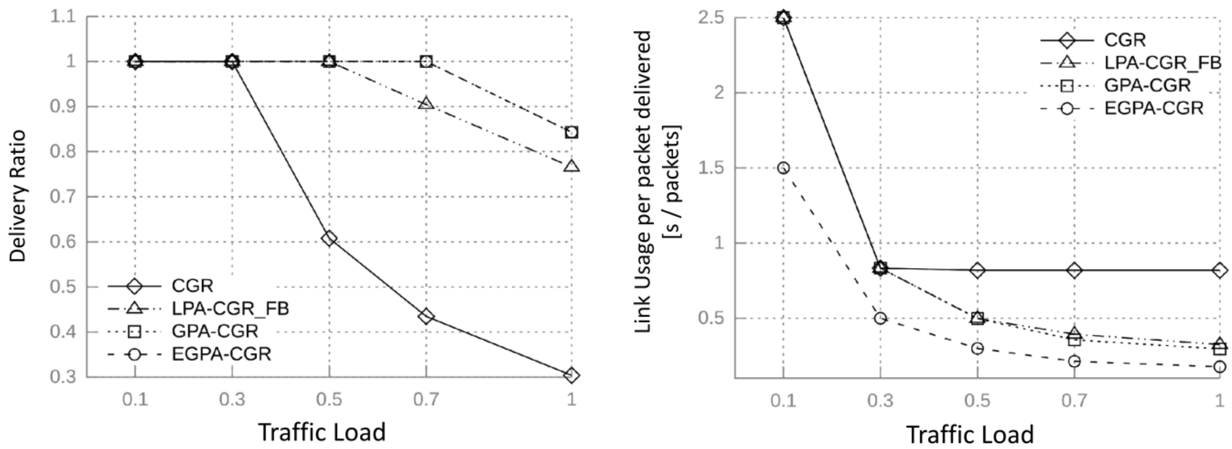

Fig. 2: Metrics comparing routing solutions with different congestion avoidance capabilities

5. Conclusion

Since future satellite networks will be subject to different conditions than terrestrial networks, new techniques will be required to avoid or mitigate congestion. In this work, different schemes that proactively avoid congestion have been compared. The results show that the greater the topological and traffic information that can be exploited, the greater the benefits in terms of delivery ratio and link utilization.

New lines of research are seeking to extend these techniques to situations where there is uncertainty regarding the topology and/or the expected traffic in the future [9]. We believe that the incorporation of Machine Learning techniques may be appropriate to achieve even more efficient proactive schemes. This would certainly reduce the need for end-to-end feedback-based schemes that introduce considerable overhead and work sub-optimally on networks for which they were not designed.

References

1. T. Ahmmed, A. Alidadi, Z. Zhang, A. U. Chaudhry and H. Yanikomeroglu, "The Digital Divide in Canada and the Role of LEO Satellites in Bridging the Gap," IEEE Magazine, 2022.
2. S. Kota, G. Giambene, L. Zhao, K. Zhao, D. Tarchi, A. Ahmed, and A. Dwivedi, "IEEE International Network Generations Roadmap (INGR) Satellite Chapter", 2022.
3. S. Kassing, D. Bhattacherjee, A. B. Águas, J. E. Saethre and A. Singla, "Exploring the "Internet from space" with Hypatia," in Proceedings of the ACM Internet Measurement Conference (IMC '20), Virtual, 2020.
4. A. Silva, S. Burleigh, C. Hirata, K. Obraczka, "A survey on congestion control for delay and disruption tolerant networks,"Ad Hoc Networks, 2015.
5. P. Madoery, J. Fraire, J. Finochietto, "Congestion management techniques for disruption-tolerant satellite networks," International Journal of Satellite Communications and Networking," 2018.

6. J. Fraire, O. De Jonckère, S. Burleigh, "Routing in the Space Internet: A contact graph routing tutorial," Journal of Network and Computer Applications, 2021.
7. "NASA: A-Train," http://atrain.nasa.gov/ (accessed November 15, 2022).
8. J. Fraire, P. Madoery, F. Raverta, J. Finochietto and R. Velazco, "DtnSim: Bridging the Gap between Simulation and Implementation of Space-Terrestrial DTNs," 2017 6th International Conference on Space Mission Challenges for Information Technology (SMC-IT), 2017.
9. F. Raverta, J. Fraire, P. Madoery, R. Demasi, J. Finocchietto, P. D'Argenio, "Routing in Delay-Tolerant Networks under uncertain contact plans," Ad Hoc Networks, 2021.

Pablo G. Madoery received the B.S. degree in telecommunication engineering from the National Defense University, Argentina, in 2015, and the Ph.D. degree in engineering sciences from the National University of Cordoba, Argentina, in 2019. He is currently a postdoctoral fellow at Carleton University on a MITACS fellowship, and an assistant professor of computer science at the National University of Cordoba, Argentina. He has co-authored 10 articles in international journals and 20 in leading conferences in the area of communications protocols and routing algorithms for delay-tolerant satellite networks.

Gunes Karabulut Kurt received the B.S. degree in electronics and electrical engineering from the Bogazici University, Istanbul, Turkey, in 2000 and the M.A.Sc. and the PhD. degrees in electrical engineering from the University of Ottawa, Canada, in 2002 and 2006, respectively. She is currently an Associate Professor of Electrical Engineering at Polytechnique Montréal, Montréal, QC, Canada. She is currently serving as an associate technical editor of the IEEE Communications Magazine, an associate editor of IEEE Communication Letters, an associate editor of IEEE Wireless Communications Letters, and an area editor of IEEE Transactions on Machine Learning in Communications and Networking.

Halim Yanikomeroglu received the BSc degree in electrical and electronics engineering from the Middle East Technical University, Turkey, in 1990, and the MASc degree in electrical engineering and the PhD degree in electrical and computer engineering from the University of Toronto, in 1992 and 1998, respectively. Dr. Yanikomeroglu's primary research domain is wireless communications and networks. His group's current focus is the wireless infrastructure for the 6G and B6G era with terrestrial, aerial (UAV), stratospheric (HAPS), and satellite network elements. He has coauthored around 550 published peer-reviewed research papers including 240+ papers in 28 different IEEE journals.

Peng Hu received the Ph.D. degree in electrical engineering from Queen's University, Canada. He is currently a Research Officer with the National Research Council of Canada, and an Adjunct Professor with the University of Waterloo. His current research interests include autonomous networking, AI-enabled edge computing, and the Internet of Things systems. He has served as a member for the IEEE Sensors Standards Committee and the organizing and technical boards/committees for industry consortia and international conferences, including AllSeen Alliance, DASH7, IEEE PIMRC'17, and IEEE AINA'15. He serves as an Associate Editor for the Canadian Journal of Electrical and Computer Engineering.

Khaled Ahmed received the BSc and MSc degrees from the Faculty of Engineering, Cairo University, Cairo, Egypt, in 2010 and 2015, respectively, and the PhD degree from McMaster University, Hamilton, Canada, in 2019. After a postdoctoral fellowship at McMaster University, he joined Satellites Systems MDA in 2021 as an Optical communications system specialist, working on system level, network level, and silicon level (architecture), and performing research and development (R&D).

Guillaume Lamontagne received the BEng and MEng degrees from the École de Technologie Supérieure, Montréal, QC, Canada, in 2007 and 2009, respectively. His experience in satellite communications started through internships and research activities with the Canadian Space Agency in 2005, and the Centre national d'études spatiales, France, in 2006 and 2008. He joined MDA in 2009 and held various communication systems engineering and management positions before being appointed as the Director of Technology, Payloads in 2019. Through this role, he is leading MDA's Research and Development activities for satellite communications as well as establishing the related long-term development strategy.