A Review on AI-Enabled Congestion Control Schemes for Content Centric Networks

Arooj Masood*, Nhu-Ngoc Dao†, Hyosu Kim* Yongseok Son* Hyung Tae Lee* Jeongyeup Paek*, and Sungrae Cho*

*School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, South Korea.
†Department of Computer Science and Engineering, Sejong University, Seoul 05006, South Korea.
Email: arooj@uclab.re.kr, nndao@sejong.ac.kr, {hskimhello, sysganda, hyungtaelee, jpaek, srcho}@cau.ac.kr.

***Abstract*—Content centric networks (CCN) offer more advantages over conventional TCP/IP networks in areas like content distribution. However, congestion control functionalities in CCN present challenges such as detecting congestion, over-reducing windows for non-congested paths, and addressing fairness issues. Most existing studies employed congestion control mechanisms similar to those in TCP. In addition, the existing mechanisms were based on conventional optimization rules to adjust the rate at which Interest packets are sent to request data from downstream nodes. However, such existing mechanisms do not consider the changes in network status and caching strategy due to multi-path and multi-source transmission. Moreover, they are based on assumptions about link bandwidth. In this paper, we study the problem of congestion control in CCN and discuss its challenges. In addition, we review the existing congestion control schemes in CCN based on machine learning. Finally, we highlight the open research issues to spur further investigations.**

***Index Terms*—Content centric networks, in-network caching, congestion control**

I. Introduction

In recent years, mobile data traffic has experienced significant expansion, and this expansion is projected to persist in the forthcoming future. Accordingly, the need for applications that require minimal delays is also rising. According to the Ericsson mobile data traffic, video traffic is estimated to account for 71% of all mobile data traffic, and this has been forecasted to increase to 80% in 2028 [1]. To address this escalating demand, a new Internet architecture called content-centric networks (CCN) [2] has been introduced, which has shifted the traditional end-to-end communication model towards content-centric communication and caches content in content store of the edge device as well as intermediate routers, allowing consumers to access contents from multiple sources using content names.

In CCN, two main packet types are used: Interest packets request content, and Data packets respond to Interests. CCN employs three crucial data structures: the content store (CS), the pending Interest table (PIT), and the forwarding information base (FIB). Consumers initiate content retrieval with an Interest packet. If content is in the CS, the router sends the Data packet to consumer. However, for non-cached content, the router checks the PIT for forwarding records and the FIB for the routing guidance.

CCN provides receiver-driven pull method and the one-interest-one-data method for its multi-source transmission mode, where former involves the consumer initiating content requests through interest packets and intermediate routers providing the requested data. Whereas, the one-interest-one-data method ensures that an Interest packet retrieves just one data packet. Consequently, CCN controls data packet transmission through Interest packet transmission rate management, which bring many news challenges such as congestion detection, rate adjustment, and fairness issues. Several researches investigated these challenges of congestion control in CCN [3]–[9]. For instance, [3]. For instance, introduced hop-by-hop Interest shaping mechanism for better link utilization. [4] proposed to predict chunk locations in CCN. [5] presented a remote adaptive active queue management (RAAQM) mechanism to control bottleneck queues. [6] employed the back-pressure method to adjust Interest rate. [7] introduced a rate-based multipath-aware congestion control, [8] presented fully-distributed congestion control (FDCC) mechanism, and [9] employed a probabilistic caching strategy. Although these existing studies managed congestion compared to the rate-based congestion control mechanisms, they were limited by the strong assumptions on continuously changing network contexts in CCN.

Recently, Machine learning (ML) methods have been applied to address the challenges of congestion control in CCN [10]–[15]. For instance, [10] proposed a deep reinforcement learning (DRL) based congestion control protocol (DRL-CCP) to learn the optimal congestion control policy. [11] proposed a forwarding strategy based on Q-learning and long-short term memory (LSTM). [12] presented a multi-path congestion control (MPCC) mechanism based on an online learning method. [13] proposed to enhance the overall network performance using DRL and [14] utilized DRL to address the existing challenges in CCN. Furthermore, [15] proposed caching strategy to dynamically select the optimal caching policy based on varying network contexts.

Accordingly, this paper examines congestion control issues in CCN and reviews AI-Enabled congestion control mechanisms in CCN. We also highlight ongoing research concerns. The remaining of the paper is organized as follows: In Section II, we discuss issues regarding congestion control in

CCN and provide a detailed review on AI-enabled congestion control schemes in CCN. Then, Section III points out ongoing research issues. Finally, we summarize the paper in Section IV.

II. Congestion Control in Content Centric Networks

In this section, we provide a brief overview on the congestion can occur in CCN. Additionally, we discuss various challenges related to congestion control in CCN. Finally, we review the existing congestion control approaches proposed for the CCN.

Because of its multi-source transmission, CCN does not have a fixed end-to-end connection. However, the reasons for congestion in CCN are the same as those in TCP/IP networks. If consumers transmit an excessive number of interest packets, the returned data packets might surpass the link bandwidth and the queue length in the intermediate router buffer. This can lead to packet losses, resulting in congestion. This highlights the relevance of traditional schemes that use TCP-like congestion control algorithms in content-centric networks. Yet, relying on packet loss as a congestion indicator is costly. In CCN, consumers detect packet loss primarily through retransmission time out (RTO) due to the absence of an ACK mechanism. Moreover, since interest requests are served by different intermediate routers with different round trip times (RTTs), detecting congestion using RTT and RTO may render inaccurate detection results. Similar to inaccurate congestion detection, traditional congestion control mechanisms adjust congestion window if there is a packet loss occurrence on any of the path. For example, in its multi-source transmission mode, if a consumer consistently receives data from two separate sources and identifies a packet loss from any bottleneck links on an intermediate router (source), the congestion window is reduced by half. This approach minimizes unwarranted traffic reduction on non-congested paths, as shown in Fig. 1. Since CCN allows in-networking caching of popular content, consumers can request popular content. As a result, consumers requesting popular content can receive content with shorter delay and may occupy more bandwidth at the bottleneck links, causing fairness issues among competing flows.

A. Congestion Control Methods in CCN

In recent years several congestion control approaches have been proposed for CCN [3]–[9]. For instance, [3], introduced a hop-by-hop interest shaping mechanism for NDN. They mathematically formulated and derived an optimal shaping rate. Subsequently, they introduced an interest-shaping algorithm designed to maximize link utilization without experiencing data loss from congestion. [4] introduced a receiver-driven congestion control protocol that addresses challenges such as content chunks being retrieved from various nodes/caches by using an anticipated interests mechanism to predict chunk locations before they are served. [5] proposed a remote adaptive active queue management (RAAQM) mechanism that controls bottleneck queues along different paths. The evaluated performance of the proposed mechanism using CCN packet-level simulations considering both random and optimal route selection. The proposed congestion control mechanism demonstrated promising results in managing multipath communication efficiently over CCN. [6] introduced a congestion control mechanism that incorporates content popularity prediction to optimize network resource allocation. Furthermore, they employed the back-pressure method to adjust the Interest rate. The results showed that the proposed popularity-centric method outperformed rate-based systems, especially in metrics like cache, Interest retransmission rate, hit ratio, goodput and flow completion time. [7] introduced a rate-based, multipath-aware congestion control algorithm. The algorithm uses a multipath forwarding strategy together with a multipath feedback mechanism. This design calculates the transmission rate for each flow and offers feedback to the consumer, guiding the sending rate of interest packets. The results demonstrated that the proposed algorithm achieves higher total throughput ensuring fairness. To overcome the dynamic location of content chunks and the lack of duplicated acknowledgements, [8] introduced the cooperative and memory-efficient token bucket (CMTB) for hop-by-hop congestion control and the fully-distributed congestion control (FDCC) for consumer-driven control. Their experimental setup included both single-path and multipath scenarios. The results showed an improvement throughput compared to other methods. [9] introduced C3NDN scheme that is based on a probabilistic caching strategy for caching the popular content. They evaluated the performance of C3NDN using the ndnSIM tool. The experimental results indicate that C3NDN can reduce transmission time compared to other schemes.

Although the conventional congestion control mechanisms such as [3]–[9] managed congestion in CCN, and demonstrated relatively good performance in terms of Interest transmission rate, throughput, and flow completion time, as compared to the rate-based methods. These studies were based on assumptions on the network characteristics and did not consider the varying network status, caching capacity, consumers' request patterns, content popularity, and application types in CNN.

B. AI-Enabled Congestion Control Methods in CCN

Machine learning (ML) approaches such as reinforcement learning and deep learning have been increasingly applied to the problem of congestion control in CCN due to the limitations of conventional methods [10]–[15]. For instance, [10] introduced a congestion control scheme that aligns with the NDN's unique characteristics, such as connectionless communication, in-network caching, and content name. They proposed DRL-CCP (a DRL-based congestion control protocol) that enables consumers to autonomously learn the optimal congestion control policy from historical data. The DRL-CCP was designed to adjust the Interest sending window size on the consumer side, considering the dynamic requirements for different content types. The results demonstrated that the proposed DRL-CCP achieves better performance in managing congestion in NDN environments.

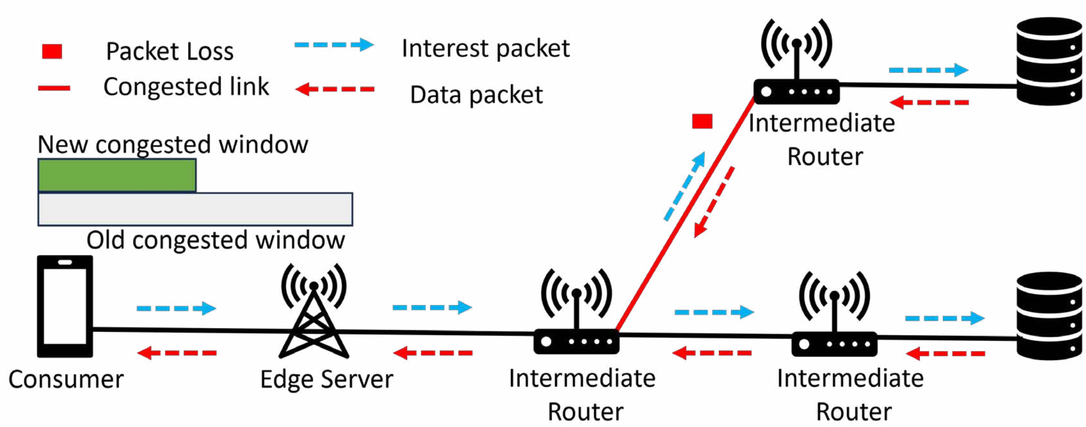

Fig. 1. Congestion window adjustment for non-congested path in CCN

[11] proposed an intelligent forwarding strategy for congestion control by integrating Q-learning and long short-term memory (LSTM). Considering the unique features of CCN, such as content-oriented communication, caching contents, multiple paths, and multiple sources, the authors addressed the congestion challenges in high-demand scenarios like video streaming. The proposed strategy is a two-phase approach. In the first phase, an LSTM model is trained to predict the pending interest table (PIT) entry rate, which can serve as an indicator of congestion. In the second phase, based on the predicted PIT entry rate, Q-learning is employed to forward data to an alternative, non-congested path. Simulation results demonstrated that their method increased the data reception rate and reduced the packet drop rate. [12] presented a multi-path congestion control (MPCC) mechanism for NDN. Considering the challenges posed by CCN's inherent multi-path and multi-source transmission, the proposed MPCC encompasses both multi-path discovery and congestion control. For discovery, they developed a unique path tag to mark each sub-path during forwarding and introduced a tag-aware strategy to manage these sub-paths. For congestion control, they considered metrics such as packet loss, bandwidth, and RTT and employed the upper confidence bound (UCB) algorithm to optimize sub-path selection to enhance network throughput. Additionally, they designed a window adaptation algorithm to prevent congestion across multiple paths. The results indicated its ability to discover all sub-paths for multi-path scenarios, and enhancing throughput while reducing transmission time. [13] researched congestion control mechanism for software-defined networking (SDN) and NDN within satellite networks based on DRL. Considering the unique characteristics of the satellite networks, the authors proposed to enhance the overall network performance by effectively managing congestion, ensuring data transmission efficiency, and maintaining network stability.

Addressing the challenges of congestion detection, window adjustment, and fairness, [14] proposed a DRL based intelligent edge-aided congestion control (IEACC) scheme for NDN. They introduced a proactive congestion detection system that leverages intermediate routers to relay precise congestion data to consumers. The IEACC approach categorizes data packets based on varying congestion levels, utilizing a lightweight clustering algorithm. This classification offers appropriate inputs for DRL to determine an optimal transmission rate. The resultsshowed that the IEACC scheme could improve the data transmission rate, maintain fairness and reduce packet loss when compared to other existing methods. [15] proposed a hybrid caching strategy based on reinforcement learning, called Cache-MAB. The proposed strategy allows routers to dynamically choose the best caching policy, considering fluctuating network scenarios. These scenarios include changes in caching capacity, user request behaviors, content popularity distribution, and types of applications. Cache-MAB aims to select the optimal policy that maximizes local performance indicators, such as cache hit rate. Simulations using ndnSIM showed that Cache-MAB effectively adapts to different network scenarios achieving near-optimal policy performance.

III. Open Research Issues

Congestion control in content-centric networks remains a dynamic research area with various unresolved challenges. Traditional TCP/IP-based congestion control mechanisms do not address the distinct requirements and characteristics of CCN. For instance, the dynamic location of content chunks, the absence of ACKs,varying RTTs, and the pull-based method of content retrieval further complicates the design of efficient congestion control schemes. While there have been significant research efforts in this area, there's a continuous quest for works that can optimize network resource allocation, cache management, and data reliability, especially in network scenarios with high packet losses. Moreover, the integration of advanced ML methods like DRL and proactive congestion detectors facilitates more intelligent and adaptive solutions. However, the practical implementation of these methods, their scalability, and their performance in real-world scenarios remain areas that require exploration.

IV. SUMMARY

With the expanding mobile data traffic over the Internet, a new Internet architecture called content centric network has been introduced. CCN has shifted the traditional end-to-end communication model towards content centric communication, presenting a response to the challenges posed by the increasing demand for on demand services. However, there exist issues such as congestion detection, congestion adjustment, and fairness. This paper discussed congestion control problems, reviewed ML-based congestion control methods, and highlighted challenges and future research directions in the field.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No.2022R1A4A5034130).

REFERENCES

[1] "Ericsson mobility report data and forcasts: Mobile data traffic outlook," Aug, 2023. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast.

[2] S. H. Ahmed, S. H. Bouk, and D. Kim, "Content-centric networks: an overview, applications and research challenges," 2016.

[3] Y. Wang, N. Rozhnova, A. Narayanan, D. Oran, and I. Rhee, "An improved hop-by-hop interest shaper for congestion control in named data networking," *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 55–60, 2013.

[4] L. Saino, C. Cocora, and G. Pavlou, "Cctcp: A scalable receiver-driven congestion control protocol for content centric networking," in *2013 IEEE international conference on communications (ICC)*. IEEE, 2013, pp. 3775–3780.

[5] G. Carofiglio, M. Gallo, L. Muscariello, and M. Papali, "Multipath congestion control in content-centric networks," in *2013 IEEE conference on computer communications workshops (INFOCOM WKSHPS)*. IEEE, 2013, pp. 363–368.

[6] H. Park, H. Jang, and T. Kwon, "Popularity-based congestion control in named data networking," in *2014 sixth international conference on ubiquitous and future networks (ICUFN)*. IEEE, 2014, pp. 166–171.

[7] S. Zhong, Y. Liu, J. Li, and K. Lei, "A rate-based multipath-aware congestion control mechanism in named data networking," in *2017 IEEE international symposium on parallel and distributed processing with applications and 2017 IEEE international conference on ubiquitous computing and communications (ISPA/IUCC)*. IEEE, 2017, pp. 174–181.

[8] A. Ndikumana, S. Ullah, K. Thar, N. H. Tran, B. J. Park, and C. S. Hong, "Novel cooperative and fully-distributed congestion control mechanism for content centric networking," *IEEE Access*, vol. 5, pp. 27 691–27 706, 2017.

[9] D. Qu, J. Wu, J. Zhang, C. Gao, H. Shen, and K. Li, "Efficient congestion control scheme based on caching strategy in ndn," *Journal of Network and Computer Applications*, vol. 216, p. 103651, 2023.

[10] D. Lan, X. Tan, J. Lv, Y. Jin, and J. Yang, "A deep reinforcement learning based congestion control mechanism for ndn," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–7.

[11] S. Ryu, I. Joe, and W. Kim, "Intelligent forwarding strategy for congestion control using q-learning and lstm in named data networking," *Mobile Information Systems*, vol. 2021, pp. 1–10, 2021.

[12] F. Wu, W. Yang, M. Sun, J. Ren, and F. Lyu, "Multi-path selection and congestion control for ndn: An online learning approach," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1977–1989, 2020.

[13] Z. Xing, H. Qi, X. Di, J. Liu, and L. Cong, "Deep reinforcement learning based congestion control mechanism for sdn and ndn in satellite networks," in *International Conference on Mobile Wireless Middleware, Operating Systems, and Applications*. Springer, 2022, pp. 13–29.

[14] J. Yang, Y. Chen, K. Xue, J. Han, J. Li, D. S. Wei, Q. Sun, and J. Lu, "Ieacc: an intelligent edge-aided congestion control scheme for named data networking with deep reinforcement learning," *IEEE Transactions on Network and Service Management*, vol. 19, no. 4, pp. 4932–4947, 2022.

[15] S. M. A. Iqbal and Asaduzzaman, "Cache-mab: A reinforcement learning-based hybrid caching scheme in named data networks," *Future Generation Computer Systems*, vol. 147, pp. 163–178, 2023.