On multiplexing data streams using trellis-coded modulation in centralized wireless networks

N. A. Yankovskii[a], *Student, orcid.org/0000-0001-5783-8304*
I. A. Pastushok[a], *PhD, Tech., Associate Professor, orcid.org/0000-0002-3296-562X, igpastushok@gmail.com*
[a]*Saint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaia St., 190000, Saint-Petersburg, Russian Federation*

Introduction: The proliferation of services and applications requiring ultra-low latency and high reliability of data transmission in communication networks leads to creating new approaches and architectures in order to ensure the simultaneous transmission of Enhanced Mobile Broadband (eMBB) and Ultra-Reliable and Low Latency Communication (URLLC) traffic. Providing efficient eMBB and URLLC multiplexing schemes with preset key performance indicators for each stream is the most challenging problem in wireless network development. **Purpose:** To provide a simultaneous transmission of eMBB and URLLC streams without reducing the user experience of eMBB services by developing a multiplexing scheme and the coherent architecture of physical (PHY) and media access control (MAC) layers in the downlink channel. **Results:** An eMBB and URLLC multiplexing scheme has been proposed, along with a coherent architecture for PHY and MAC layers, ensuring the given wireless network key performance indicators. The proposed solution performance has been estimated by simulation. The multiplexing scheme outperforms the baseline solution in Bit Error Rate and Frame Error Rate metrics. The coherent PHY and MAC layers architecture provides transmission with an arrival rate of 400 messages per millisecond and 99% message delivery probability in one millisecond. **Practical relevance:** The obtained results allow communication system developers to deploy centralized wireless networks at industrial objects.

Keywords — URLLC, eMBB, multiplexing, radio resource management, centralized wireless networks, trellis-coded modulation, LDPC.

Introduction

In just a few decades, wireless technology has undergone rapid growth from its original concept to ubiquitous penetration, which has changed our daily lives and thinking. Wireless connection has become an essential link between people and information networks. The growing density of user devices triggered an increasing demand for higher capacity and network reliability. The constant increase in traffic leads to congestion of base stations and a decrease in service level quality.

An attractive solution to this problem is the multiplexing of data streams in the downlink. A substantial amount of in-depth research has been dedicated to this topic.

In [1], authors consider various models for the Enhanced Mobile Broadband (eMBB) rate loss associated with Ultra-Reliable and Low Latency Communication (URLLC) superposition/puncturing, for which we characterize the associated feasible throughput regions and online joint scheduling algorithms. The first model considered by the authors is the linear model. When the rate loss to eMBB is directly proportional to the fraction of superposed/punctured mini-slot. The second model considered was the convex model, where the rate loss can be modeled through a convex function. And the last model considered was the threshold model where eMBB traffic is unaffected by puncturing until a threshold. Beyond this threshold, it suffers complete throughput loss.

There are also several papers describing a specific scheme for the coexistence of multiple data streams. So, in [2], the authors introduced an approach for coexisting URLLC [3] and eMBB [4] traffic in the same radio resource for enabling 5G wireless systems. They have expressed the coexisting dilemma as a maximizing problem of the minimum expected achieved rate value of eMBB user equipment (UEs) meanwhile attending the URLLC traffic.

Also, they presented a heuristic algorithm for the efficient scheduling of resource blocks among eMBB UEs. In [5], the authors considered approaches to data multiplexing based on machine learning. In their work, they proposed an optimization-aided deep reinforcement learning-based algorithm, which proposed to distribute the incoming URLLC traffic among eMBB users intelligently. In [6], the authors consider the optimization problem of maximizing the transmission rate of eMBB traffic, subject to URLLC requirements. To study the impact of puncturing eMBB resources to accommodate URLLC transmission, the authors in [7] investigated the problem of joint planning of eMBB and

URLLC data transmission according to linear, convex, and threshold velocity stall models by eMBB associated with the drilling of the eMBB resource. In [8], a risk-sensitive approach was introduced to mitigate the risk of puncturing into eMBB resources. A resource allocation planner was proposed in [9], where the formulated problem considered the overhead associated with URLLC load segmentation while maximizing speed utility. In [10], a null space-based spatial perforation scheduler for joint URLLC / eMBB traffic has been proposed. The authors in [11] formulated a URLLC traffic allocation problem by adopting an overlay or perforation scheme. In practice, when the URLLC service is started in the middle of the eMBB transport block, part of the eMBB symbols are replaced and/or overlapped with the symbols in the URLLC packet. As a result, the reception quality of eMBB services can be significantly reduced.

In [12], the authors also study the orthogonal and non-orthogonal slicing of radio resources for eMBB and URLLC using a maximum matching diversity (MMD) algorithm to locate frequency channels of eMBB users. In [13], the authors adopted a time/frequency resource block approach to address the problem of maximizing the sum rate subject to latency and cutoff isolation constraints while ensuring the reliability requirements using adaptive modulation coding. In [14], the authors studied a multi-cell scenario with a single-cell base stations for an Ultra-Narrow Band and Low Power Wide Area Network. Article [15] analyzes the use of non-orthogonal multiple accesses (NOMA) for different URLLC devices. To achieve this, the authors propose a NOMA sharing approach, successive interference cancellation, and frequency diversity as a solution to increase the number of URLLC devices that can be connected to the same base station. In [16], a new class of NOMA has been proposed, namely bits similarity NOMA. It has been shown that without a perfect successive interference cancellation, bit similarity NOMA can achieve better efficiency between users than traditional NOMA techniques.

Another approach to multiplexing URLLC and eMBB traffic is Trellis- and Network-coded modulation (TC-NCM) [17]. Based the Ungerboeck's scheme [18] and the general type of coset coding advocated by Goldsmith in [19] and Chapter 8 of [20], the authors propose adaptive TC-NCM structure, where the transmitter adapts the coding rate and modulation mode according to the channel estimates fed back via feedback channels. The main disadvantage of this method is the lack of encoding of the URLLC stream, which negatively affects the reliability and transmission rate of URLLC messages. This work's main task is to propose a Trellis-Coded Modulation (TCM) scheme for multiplexing streams and define the loss function for it. The remainder of this paper is organized as follows. We commence by describing the system model along with a system of assumptions. We then conceive our generic structure of the multiplexing scheme, where the motivation, the transmitter design, and the data flow are detailed. Finally, we present some modeling results and conclusions.

Baseline system model

In this article, we discuss the development of the model proposed in [21] generalized to the physical layer of wireless centralized systems. As before, the baseline scenario will be the case when only 1 data stream is transmitted. However, in contrast to the previous work [21], we will consider the Viterbi decoder's quantized output — the general scheme for basic scenarios presented in Fig. 1.

In the case of a multiplexing scenario, the network contains a base station, as well as users receiving eMBB and URLLC traffic from the base station, respectively. As in the previous article, the simultaneous transmission is considered data for different recipients from one sender. Each data stream has its own scheduler. The scheduler for eMBB traffic uses time division multiplexing (TDM/TDMA), and the scheduler for URLLC uses (OFDM/OFDMA) (Fig. 2).
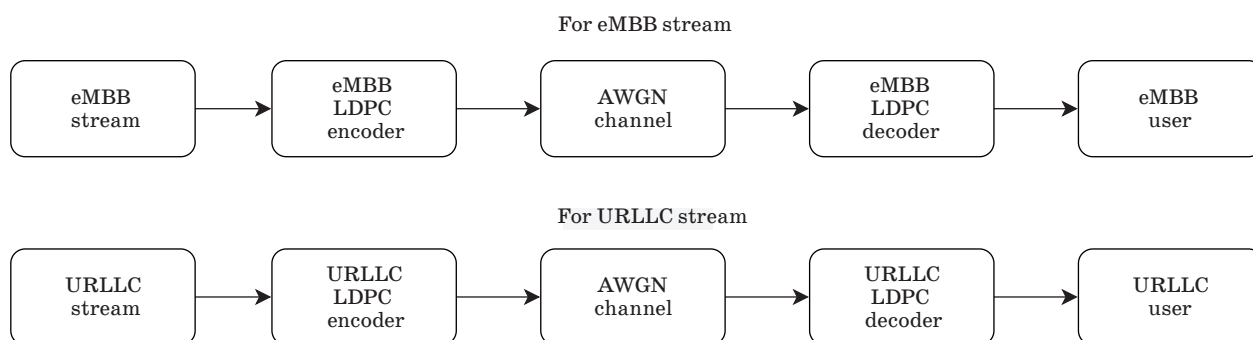
■ *Fig. 1.* Baseline scenario
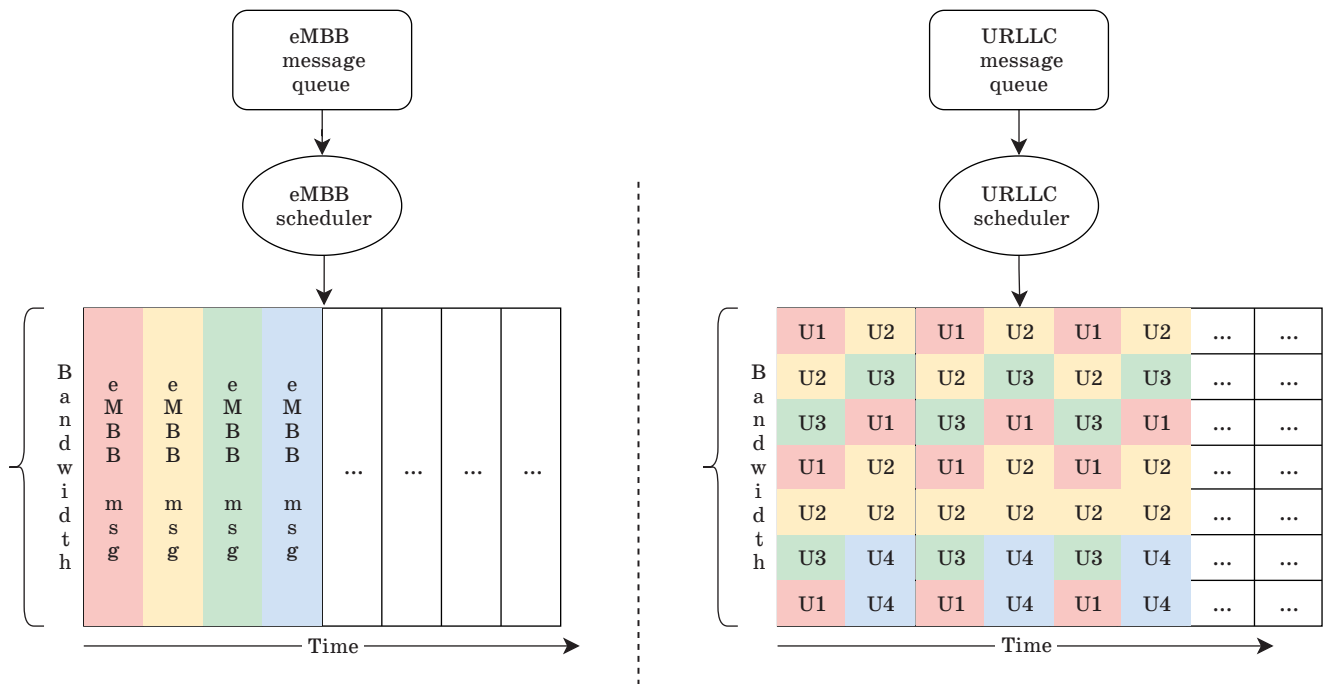
■ *Fig. 2.* Resource scheduling for different data streams
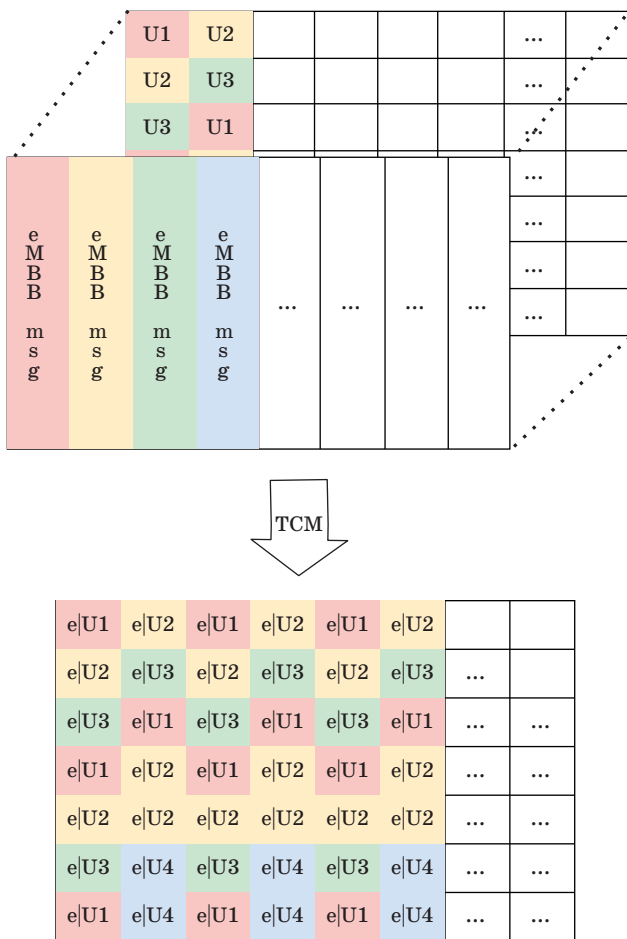
■ *Fig. 3.* Scheduling resources after using multiplexing

After the allocation of the radio resource by the schedulers, a multiplexing scheme is used, after which the information bits of different streams are located in the combined resource block for further transmission over the channel (Fig. 3).

Assumptions

Having outlined the transmission model, next we list all of our operating assumptions used throughout this paper.

1. Channel state information is always available on the sender. Assume that the feedback path does not introduce any errors, which can be approximately satisfied, provided that sufficiently powerful error correction and detection codes are used on the feedback path.

2. We are considering a channel with additive white Gaussian noise.

3. We consider the downlink in TDMA mode for eMBB data stream and OFDMA for URLLC data stream.

4. We know on the receiving side of the URLLC traffic transmission position for the user.

5. EMBB users are not aware of the existence of URLLC traffic and do not decode it.

Multiplexing algorithm

It is known that the Ungerboeck scheme [18] combines encoding and modulation by expanding

the Euclidean distance between codewords and absorbs parity bits without bandwidth expansion by doubling the number of points in the constellation due to increasing the number of bits / symbols by one. This design jointly optimizes both channel encoding and modulation, hence again, resulting in significant encoding gains without any bandwidth expansion. Based on these ideas, our adaptive stream multiplexing scheme was developed.

Generic multiplexing scheme

At some point in time, an eMBB message and several URLLC messages appear on this station. Then the algorithm for multiplexing and transmitting messages is as follows (Fig. 4):

1. A checksum is added to all messages.

2. Messages are fed to the LDPC encoding unit. For eMBB traffic the code rate is 8/9, and for URLLC traffic it is 1/2. This choice is due to the fact that URLLC messages are shorter and more demanding on the reliability of transmission.

3. After LDPC encoding, the messages are combined into one using an interleaver.

4. The general message is fed to the input of the TCM, which selects the modulation dimension required to transmit $(l + r)$ symbols, where l — message length; r — redundancy of the applied convolutional code.

5. Then the message is transmitted over the communication channel and enters the input of the decoder, which is a soft output Viterbi algorithm.

6. The message is quantized according to the following rule: the most significant $\log_2(Q)$ bits of the Log-Likelihood Ratio are stored, where Q is the number of quantization levels.

7. The codewords of each stream passes the deinterleaver and LDPC decoder.

8. The checksum is checked, and a decision is made on the correctness of the received message.

Key performance indicators of the system

To assess the quality of the proposed multiplexing algorithm, it is necessary to introduce indicators of efficiency. This article discusses the following key performance parameters:

1) frame error rate (FER) for URLLC and eMBB streams;

2) bit error rate (BER) for eMBB stream;

3) channel capacity for URLLC stream;

4) complexity of separating eMBB and URLLC streams.

To evaluate the first two performance criteria, we introduce the following notation A_{ber} and B_{ber} — functions for the considered scenarios that return the SNR (signal-to-noise ratio) value to achieve the required BER value. A_{fer} and B_{fer} — return SNR value to achieve the required FER respectively. Thus, we consider that system B is not inferior to system A if:
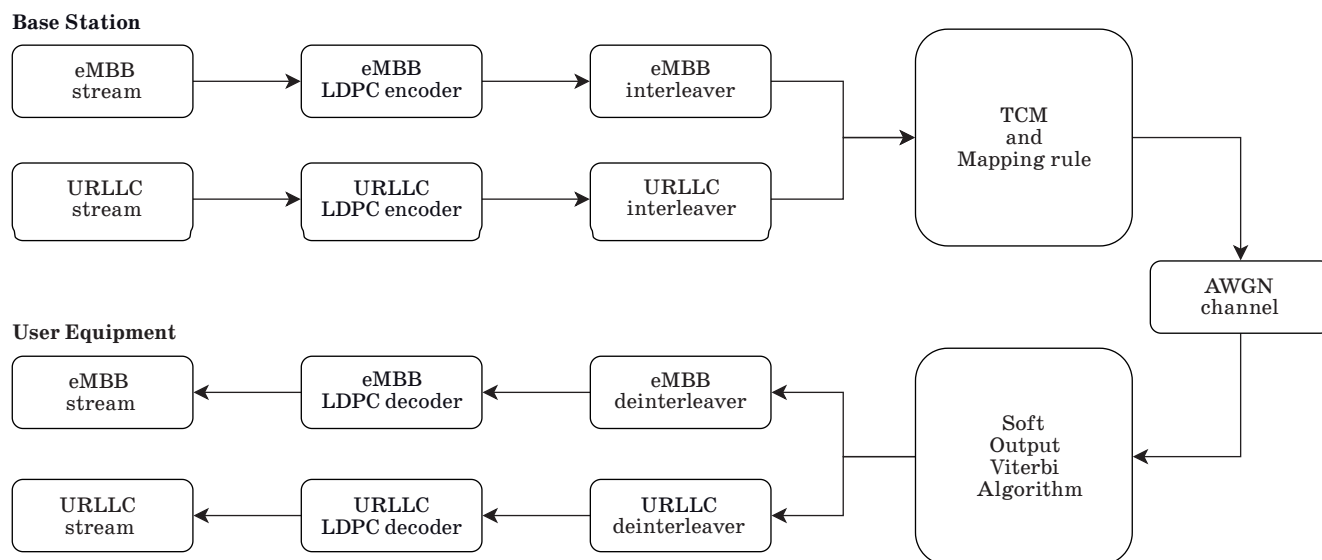
$$B_{ber}(10^{-7}) <= A_{fer}(10^{-7});$$
$$B_{fer}(10^{-4}) <= A_{fer}(10^{-4});$$

frequency bands of A and B are the same.

The channel capacity here means the ability to transmit all URLLC messages in one time slot. The following algorithm was used to calculate the channel capacity for the URLLC stream:

1. A set of messages is generated according to the Poisson distribution.

■ *Fig. 4*. General scheme of the multiplexing algorithm

2. The generated messages occupy slots in the current time slot.

3. Messages that did not get free resource blocks or were incompletely allocated are discarded and replenished the message buffer that was refused transmission.

4. Based on the number of discarded messages and the total number, the probability is calculated that the message will not be transmitted in 1 time slot.

Stream splitting algorithm consists of two parts: Soft-Output Viterbi Algorithm (SOVA) and deinterleaver. The time complexity of the Viterbi algorithm can be expressed as $O(NS^2)$, where N is a length of message in bits, and S is a number of states in a hidden Markov model.

The complexity of deinterleaving is equal to the size of the message, so the overall complexity of streaming can be considered equal to the complexity of decoding using the Viterbi algorithm.

System parameters

The 5G-NR standard implies the use of a different number of templates with different network parameters (Table 1).

■ *Table 1*. 5G-NR numerology

Parameter	Numerology values				
	0	1	2	3	4
Subcarrier width, kHz	15	30	60	120	240
Num of slot in subframe	1	2	4	8	16
Slot duration, ms	1	0.5	0.25	0.125	0.0625

Each template allows you to adaptively configure the physical layers of the system. Such templates are called numerology, and, in this paper, we use the numerology parameters under number 3 (see Table 1). The structure of the frame is presented in Fig. 5. This article considers the mm-Wave scenario with 50 MHz cell and numerology 3. All-time divided by subframes 1ms duration each. By numerology 3, each subframe is divided into eight slots, 0.125 ms each. Thus, we can calculate the number of available OFDM symbols for multiplexing in each slot and subframe.

LDPC codes in the 5G-NR standard also have many different parameters. The parameters used in our system are shown in Table 2.
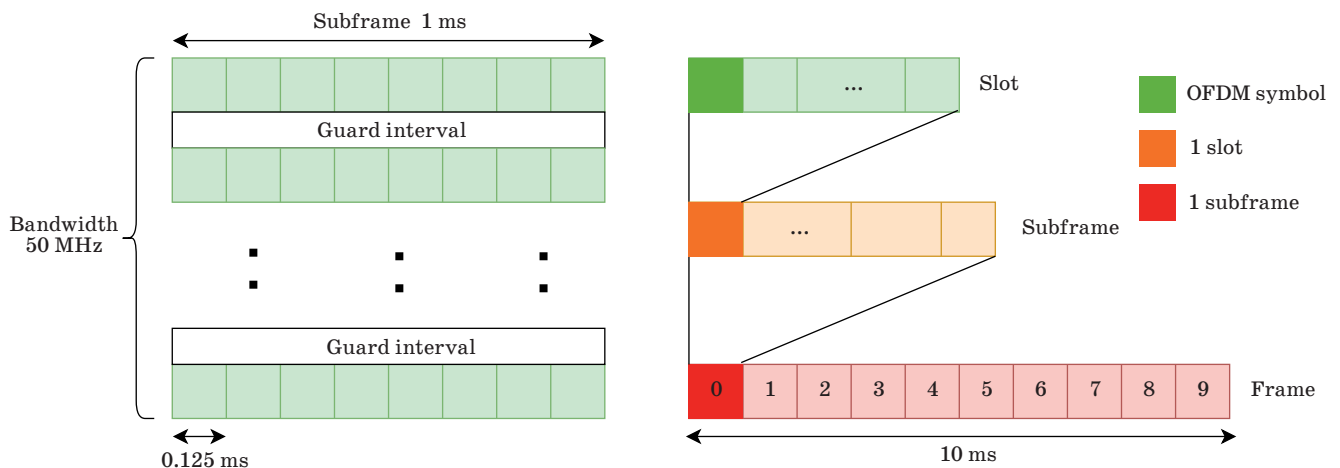
The main parameter of TCM is the use of convolutional polynomial coding. In our work, we use the polynomial presented by William G. Chambers [10] (Fig. 6), since it provides the maximal possible free

■ *Table 2*. LDPC parameters

Stream	Code rate	Reduncity level	nlayers
eMBB	8/9	0	3
URLLC	1/2	0	3

■ *Table 3*. TCM parameters

Octets	Memory	d_f	Code rate	Modulation	Mapping rule
117 155	6	10	1/2	QAM-16	Gray mapping

■ *Fig. 5*. 5G-NR Numerology 3 frame structure

G_0 — 1 0 0 1 1 1 1 117 (octal)

X →

G_1 — 1 1 0 1 1 0 1 155 (octal)

■ *Fig. 6.* Convolutional code

■ *Table 4.* Coding gain

Stream	$Q = 2$	$Q = 4$	$Q = 16$	$Q = 256$
eMBB, FER = 10^{-2}	2.5	0.8	0.7	0.3
eMBB, FER = 10^{-4}	2.6	0.5	0.5	0.3
URLLC, FER = 10^{-5}	–	–	–	–0.5
URLLC, FER = 10^{-6}	–	–	–	0.5

Table 4 shows the gain in dB when using multiplexing for different quantization levels.

URLLC stream capacity

Since the URLLC stream must fulfill the requirements of immediate transmission, we will assume that if the message was not transmitted per 1ms subframe, then it loses its relevance for the end user. Fig. 9 shows the probability of non-transfer URLLC messages with a length of 100 bits per 1 ms subframe for the considered multiplexing scenario (see simulation details in [21]). We can conclude that the probability of sending a URLLC message in one mini-slot more than 99% for an incoming rate of up to 400 messages per slot.

Thus, the proposed downlink multiplexing scheme using TCM allows to obtain a lower error probability in the channel for eMBB traffic, as well as to preserve the key parameters of the efficiency applied to URLLC traffic. The proposed scheme
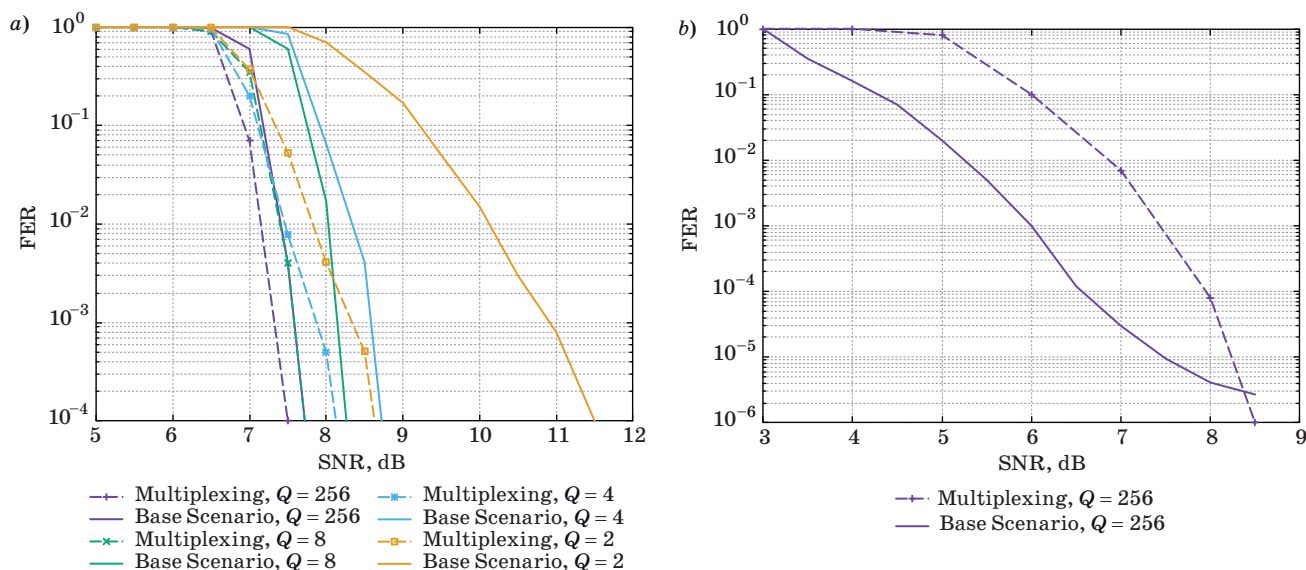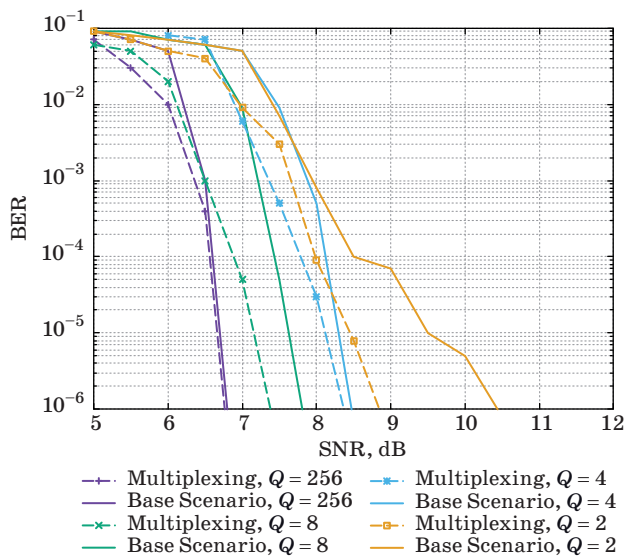
distance (d_f) and so the maximal asymptotic coding gain. The parameters of the selected polynomial are shown in Table 3.

Simulation results

Using the parameters described above, the FER versus SNR plots were obtained using simulation for both data streams (Fig. 7, *a* and *b*).

It can be seen from the figures above that due to the complication of the decoding procedure when multiplexing streams, the FER values for all quantization levels exceed the FER values for the baseline scenario.

Let us consider the efficiency of the multiplexing algorithm in terms of BER for the eMBB stream. Fig. 8 shows a plot of BER versus SNR for different quantization levels.

■ *Fig. 7.* FER versus SNR for eMBB (*a*) and URLLC (*b*) stream

■ *Fig. 8.* BER versus SNR for eMBB stream, Q is a number of quantization levels
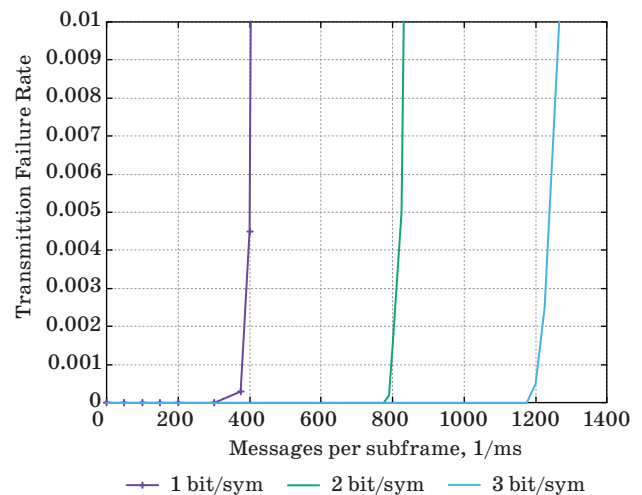

■ *Fig. 9.* Denial of service rate

can be classified as a threshold model by **[1]**, and all known results can be applied to ours.

Conclusion

This article proposed an algorithm for multiplexing data streams in the downlink communication of centralized wireless networks. This algorithm fits the threshold model proposed by Veciana G. Then, a general scheme based on TCM was considered and applied to 5G-NR Numerology 3 networks. Our simulation results show that the proposed algorithm achieves better BER and FER performance for the eMBB stream compared to the baseline scenario without multiplexing. We also presented graphs of the channel throughput for the URLLC stream, from which it can be seen that this scheme allows providing the probability of sending a URLLC message in one mini-slot more than 99% for an incoming rate of up to 400 messages per slot. For promising future research, an attractive direction is the study of achievable improvement in constellation formation, improving system performance.

Financial support

References

1. Anand A., Veciana G., and Shakkottai S. Joint scheduling of URLLC and eMBB traffic in 5G wireless networks. *IEEE/ACM Transactions on Networking*, 2020, vol. 28, pp. 477–490.
2. Bairagi K., Munir M. S., Alsenwi M., Tran N. H., Alshamrani S. S., Masud M., Han Z., and Hong C. S. *Coexistence mechanism between eMBB and uRLLC in 5G wireless networks*. 2020. Available at: https://arxiv.org/pdf/2003.04551.pdf (accessed 20 November 2020).
3. ITU-R M.2083-0. IMT Vision — framework and overall objectives of the future development of IMT for 2020 and beyond. ITU-R, 2015. 21 p.
4. 3GPP TS 22.261 16.4.0. Service requirements for next generation new services and markets. 3GPP, 2017. 55 p.
5. Alsenwi M., Tran N. H., Bennis M., Pandey S. R., Bairagi A. K., and Hong C. S. *Intelligent resource slicing for eMBB and URLLC coexistence in 5G and beyond: A deep reinforcement learning based approach*. Available at: https://arxiv.org/pdf/2003.07651.pdf (accessed 14 December 2020).
6. Alsenwi M., Hong C. S. *Resource scheduling of URLLC/eMBB traffics in 5G new radio: A punctured scheduling approach*. Avaliable at: http://networking.khu.ac.kr/xe/Gallery/entry/document_srl/141/page/layouts/net/publications/data/KCC2018/14.Madyan.pdf (accessed 14 December 2020). doi:10.1109/INFOCOM.2018.8486430.
7. Anand A., Veciana G., and Shakkottai S. Joint scheduling of URLLC and eMBB traffic in 5G wireless networks. *IEEE INFOCOM 2018 — IEEE Conference on Computer Communications*, April 2018, pp. 1970–1978.
8. Alsenwi M., Tran N. H., Bennis M., Kumar Bairagi A., and Hong C. S. eMBB-URLLC resource slicing: A risk-sensitive approach. *IEEE Commun. Lett.*, 2019, vol. 23, no. 4, pp. 740–743.

9. Karimi A., Pedersen K. I., Mahmood N. H., Pocovi G., and Mogensen P. Efficient low complexity packet scheduling algorithm for mixed URLLC and eMBB traffic in 5G. *IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, April 2019, pp. 1–6.

10. Khalifa N. B., Angilella V., Assaad M., and Debbah M. Low-complexity channel allocation scheme for URLLC traffic. *IEEE Trans. on Commun.*, 2020. doi:10.1109/TCOMM.2020.3022008

11. Manzoor A., Kazmi S. M. A., Pandey S. R., and Hong C. S. Contract-based scheduling of URLLC packets in incumbent EMBB traffic. *IEEE Access*, 2020, vol. 8, pp. 167516–167526. doi:10.1109/ACCESS.2020.3023128

12. Santos E. J., Souza R. D., Rebelatto J. L., and Alves H. Network slicing for URLLC and eMBB with max-matching diversity channel allocation. *IEEE Commun. Lett.*, 2020, vol. 24, no. 3, pp. 658–661.

13. Korrai P. K., Lagunas E., Sharma S. K., Chatzinotas S., and Ottersten B. Slicing based resource allocation for multiplexing of eMBB and URLLC services in 5G wireless networks. *IEEE CAMAD*, Sep. 2019, pp. 1–5.

14. Mo Y., Goursaud C., and Gorce J. Uplink multiple base stations diversity for UNB based IoT networks. *IEEE Conference on Antenna Measurements Applications (CAMA)*, 2018, pp. 1–4.

15. Kassab R., Simeone O., Popovski P. and Islam T. Non-orthogonal multiplexing of ultra-reliable and broadband services in fog-radio architectures. *IEEE Access*, 2019, vol. 7, pp. 13035–13049. doi: 10.1109/ACCESS.2019.2893128

16. Chraiti M., Ghrayeb A., and Assi C., A NOMA scheme exploiting partial similarity among users bit sequences. *IEEE Trans. on Commun.*, 2018, vol. 66, no. 10, pp. 4923–4935.

17. Yang Y., Chen W., Li O., Ke K., and Hanzo L. Trellis- and network-coded modulation for decode-and-forward two-way relaying over time-varying channels. *IEEE Trans. Veh. Technol.*, 2017, vol. 66, no. 6, pp. 4845– 4858. doi:10.1109/TVT.2016.2615656

18. Ungerboeck G. Channel coding with multilevel/phase signals. *IEEE Trans. Inf. Theory.*, 1982, vol. 28, no. 1, pp. 55–67. doi:10.1109/TIT.1982.1056454

19. Goldsmith J., Chua S. G. Adaptive coded modulation for fading channels. *IEEE Trans. Commun.*, 1998, vol. 46, no. 5, pp. 595–602. doi: 10.1109/ICC.1997.595036

20. Goldsmith J. *Wireless communications.* Cambridge, UK, Cambridge Univ. Press, 2005. 644 p. doi:10.1017/CBO9780511841224

21. Pastushok I. A., Boikov N. A., and Yankovskii N. A. Bit stream multiplexing in 5G networks. *Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, Saint-Petersburg, Russia, 2020, pp. 1–4. doi:10.1109/WECONF48837.2020.9131511

О мультиплексировании потоков данных с использованием решетчатого кодирования в централизованных беспроводных сетях

Н. А. Янковский[а], студент, orcid.org/0000-0001-5783-8304
И. А. Пастушок[а], канд. техн. наук, доцент, orcid.org/0000-0002-3296-562X, igpastushok@gmail.com
[а]Санкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: распространение сервисов и приложений, требующих сверхнизкой задержки и высокой надежности передачи данных в сетях связи, приводит к необходимости создания новых подходов и архитектур для обеспечения одновременной передачи разнородного трафика улучшенной широкополосной связи (eMBB) и сверхнадежной связи с низкими задержками (URLLC). Одной из самых актуальных задач в области разработки стандартов беспроводной связи является обеспечение мультиплексирования потоков eMBB и URLLC с требуемыми показателями производительности передачи каждого потока. **Цель:** обеспечить одновременную передачу потоков eMBB и URLLC без потерь пользовательского опыта сервисов eMBB путем создания метода мультиплексирования потоков данных на основе решетчатого кодирования и модуляции сигнала, а также соответствующую настоящему методу архитектуру физического и канального уровней беспроводных централизованных сетей связи. **Результат:** предложены метод мультиплексирования потоков данных eMBB и URLLC в нисходящем канале связи, а также согласованная с ним архитектура физического и канального уровней сетей, позволяющие обеспечить заданные требования функционирования беспроводной сети. Оценка эффективности представленного решения путем имитационного моделирования дает возможность сделать следующие утверждения. Разработанный метод мультиплексирования обеспечивает лучшие значения вероятностей ошибки на бит и на кодовое слово в сравнении с опорным сценарием на физическом уровне. Предложенная архитектура канального уровня позволяет обеспечить передачу потока URLLC с интенсивностью 400 сообщений в миллисекунду с вероятностью доставки сообщения, равной 99%, в течение одной миллисекунды. **Практическая значимость:** полученные результаты помогут разработчикам систем связи планировать развертывания беспроводных централизованных сетей в промышленности.

Ключевые слова — URLLC, eMBB, мультиплексирование, планирование, беспроводные централизованные сети, решетчатое кодирование и модуляция, LDPC.