Service-oriented model for handling mMTC subscribers' traffic in a 5G cluster

Viacheslav Kovtun [1,*,†] and Oksana Kovtun [2,†]

[1] Institute of Theoretical and Applied Informatics Polish Academy of Sciences, Bałtycka Str., 5, Gliwice, 44-100, Poland
[2] Vasyl' Stus Donetsk National University, 600-richchya Str., 21, Vinnytsia, 21000, Ukraine

Abstract
The article addresses the research task of rational organization of the data collection process from multiple sensor networks with a finite number of terminals located in the 5G cluster coverage. The proposed solution is based on a Markov model of the controlled process of uplink transfer of segmented informational messages from a finite set of terminals belonging to different sensor networks deployed in the coverage area of the target 5G cluster. The segregation of sensor networks at the model level is ensured by their service orientation in dedicated virtual network segments organized in the information environment of the 5G base station. Moreover, the model allows for establishing priorities for services characteristic of each sensor network. The specificity of mMTC traffic is taken into account by the base station, which accepts an incoming request only if there is a minimum guaranteed amount of available communication resources necessary for its processing. It is assumed that an unaccepted request may wait for processing in a buffer of finite capacity, which it leaves either upon acceptance for service or upon expiration of the assigned waiting time. The model serves as the basis for formalizing the metrics of quality indicators, including the probability of losing an incoming request due to system overload, indicators of the average duration of a request's stay in the system, and indicators of the average number of requests in the system. To demonstrate the functionality of the proposed mathematical framework for the target service-oriented traffic handling system of mMTC subscribers in a 5G cluster, dependencies of quality metric indicators on the intensity of incoming requests to the base station were calculated. The obtained results allowed for the estimation of the volume of communication resources that can be reasonably predicted for the efficient operation of the target 5G cluster with mMTC traffic.

Keywords
5G Cluster, Sensor Network, mMTC Traffic, Communication Resource Handling, Markov Model, Quality Metrics.

Introduction

In the list of technologies proposed by the 5G/IMT-2020 standards specification [1], the support for massive Machine Type Communication (mMTC) holds a prominent position, and this is not coincidental. The implementation of Industry 4.0 leads to a continuous expansion of sensor network coverage areas, connecting them to the cloud via wired communication is neither technologically nor economically feasible, and often simply impossible. The mMTC technology is specifically designed to address a plethora of issues related to organizing informational interactions around sensor networks. However, any system operates predictably and efficiently only if its structure is transparent and optimal in the context of its intended purpose. Therefore, the research task of the rational organization of the data collection process from multiple sensor networks with a finite number of terminals, located in the vicinity of a 5G cluster, is relevant and of great practical significance.

IntelITSIS'2024: 5th International Workshop on Intelligent Information Technologies and Systems of Information Security, March 28, 2024, Khmelnytskyi, Ukraine
* Corresponding author.
† These authors contributed equally
vkovtun@iitis.pl (V. Kovtun); o.kovtun@donnu.edu.ua (O. Kovtun)
0000-0002-7624-7072 (V. Kovtun); 0000-0002-9139-8987 (O. Kovtun)

Numerous research papers have delved into the exploration and examination of radio resource management, control and allocation techniques within 5G networks, as evidenced by a collection of studies [2-4]. The intricate aspects of radio resource control and Quality of Service (QoS) provisions for 5G network slices have been introduced and scrutinized in various research endeavours [5, 6]. Recent works have also addressed the introduction and discussion of resource allocation strategies and challenges pertinent to network slicing [7-9].

The investigation of slice isolation's end-to-end behaviour, particularly from a security standpoint, is the focus of [10]. This study classifies isolations such as traffic, processing, bandwidth or storage, while also addressing challenges associated with the contemporary trends in 5G slice isolation. The authors emphasize that isolation stands out as the most pivotal characteristic of network slicing.

[18] introduces an effective and secure service-oriented authentication framework designed to facilitate network slicing within the 5G-powered Internet of Things network. The framework aims to provide a robust solution to authentication concerns.

In [11], a flexible network slicing framework is unveiled, featuring a regional orchestrator tasked with coordinating workload dispersion among nearby fog nodes.. This orchestrator enables the dynamic adjustment of allocation of resources to each slice determined by service requests and energy accessibility..

The utilization of queuing models to address resource allocation challenges in the coexistence of various services with diverse QoS requirements within 4G/5G networks is explored in [12-14]. Specifically, the co-occurrence of machine-to-machine (M2M) communications in 4G networks is investigated in [15].

In [16-18], a queuing model is employed to analyze the repercussions of the Co-occurrence between M2M communication within a New Radio (NR) system. This study aims to understand the impacts of simultaneous M2M communication on network dynamics.

Meanwhile, [19, 20] proposes a resource-sharing approach for Machine-to-Machine (M2M) traffic within a time-regulated scheduling scheme within NR networks. This approach is designed to efficiently manage and allocate resources between machine-based communications, contributing to enhanced network performance.

Taking into account the identified constraints characteristic of the aforementioned closely related studies, let's formulate the object, subject, aim and tasks of our research.

The object of the research is the controlled process of uplink transfer of segmented informational messages from a finite set of terminals belonging to different sensor networks deployed in the coverage area of the target 5G cluster.

The subject of the research is the elements of queuing theory, Markov chains, and functional analysis.

The aim of the research is to streamline approaches to the analytical assessment of the quality of the service-oriented, controlled process of handling mMTC subscriber traffic in a 5G cluster.

Tasks of the research include:

- Parameterizing the research object with the formulation of the optimization problem concerning a metric such as the volume of communication resources available to the 5G base station for handling incoming mMTC traffic;
- Formalizing a Markov service-oriented model for handling mMTC subscriber traffic in a 5G cluster;
- Formalizing quality metric indicators for evaluating the instance of a 5G cluster with mMTC traffic from multiple sensor networks;
- Analyzing empirical results obtained during the demonstration of the proposed mathematical framework's functionality.

Models and methods

Research Statement

The focus of our research is on the 5G cluster, whose base station is capable of distributing communication resources in the amount of V units. Within the coverage area of the investigated cluster, mobile network operators potentially can offer subscribers specific services, the list of which is generalized by the set $\mathrm{S}=\left\{\overline{1,S}\right\}$. In this context, the k-th operator, $k\in\mathrm{K}=\left\{\overline{1,K}\right\}$, provides its subscribers with a list of services, generalized by the subset $S_k\subseteq\mathrm{S}$. In turn, supporting the s-th service, $s\in\mathrm{S}$, in an active state requires a minimum amount of communication resources in the volume of v_s units. The number of subscribers willing to use the s-th service of the k-th operator will be denoted by the parameter N_{ks}. Accordingly, the total number of subscribers registered in the investigated 5G cluster will be determined as $N=\sum_{k=1}^{K}N_k=\sum_{k=1}^{K}\sum_{s\in\mathrm{S}_k}N_{ks}$.

Taking into account that $v_s=\mathrm{const}\quad\forall n\in\mathrm{N}=\left\{\overline{1,N}\right\}$, the process of handling the distribution of V units of communication resources among the elements of the tuple $\left\langle\mathrm{S},\mathrm{K},\mathrm{N}\right\rangle$ can be formulated as an optimization problem with the objective function

$$z\left(\mathrm{V}\right)=\sum_{k=1}^{K}\sum_{s\in\mathrm{S}_k}p_{ks}V_{ks}\rightarrow\max \tag{1}$$

and a system of constraints

$$\begin{aligned}&\sum_{k=1}^{K}\sum_{s\in\mathrm{S}_k}p_{ks}V_{ks}\le V,\\&0\le V_{ks}\le V,\end{aligned} \tag{2}$$

where the parameter V represents the matrix form of the distribution, such as $\mathrm{V}=\left(V_{ks}\right)_{s\in S_k,k\subseteq K}$, $V_{ks}>0\ \ \forall s\subseteq S_k$, $V_{ks}=0\ \ \forall s\not\subseteq S_k$; the parameter p_{ks} represents the priority of the s-th service of the k-th operator, $s\in\mathrm{S}$, $k\in\mathrm{K}$, $0\le p_{ks}\le1$; the parameter V_{ks} represents the volume of communication resources that the base station allocates to support the s-th service of the k-th operator (respectively, $V_k=\sum_{s\in S_k}V_{ks}$ is the volume of communication resources that the base station allocates to support the entire service package of the k-th operator).

The optimization problem (1), (2) is formulated considering that the investigated 5G cluster is oriented towards supporting mMTC technology. Therefore, during the problem formulation stage, we only took into account the minimum amount of communication resources v_s necessary to support the s-th service, $s\in\mathrm{S}$. For example, if we were describing eMBB (enhanced Mobile Broadband) technology, we would have needed to consider both the minimum amount of communication resources $v_s^{\max}$ and the maximum volume of communication resources $v_s^{\max}$, which would be rational to anticipate for supporting the s-th service.

Service-oriented model for handling mMTC subscribers' traffic in a 5G cluster

Let's formulate a service-oriented model for handling mMTC traffic in a 5G cluster. In doing so, we will take into account the capabilities of the 5G technology's Network Slicing feature. The implementation of this technology in the investigated 5G cluster allows the distribution of communication resources at the base station among independent virtual network segments, each assigned to respective mobile network operators. Within the allocated virtual network segment, an operator ensures the support of a declared range of proprietary services with guaranteed data transfer speeds.

We will formulate a model of the investigated process based on the queuing systems theory. Suppose N subscribers direct their requests to the base station to activate available services in the portfolios of operators, each of which is associated with a corresponding virtual network segment. It is assumed that a subscriber cannot submit a new request until receiving a response regarding the recently submitted request (this restriction is introduced to prevent potential DoS attacks). The base station has V units of communication resources available to handle subscribers' requests. The incoming requests to the base station will be characterized by the arrival intensity $\eta_n > 0$, $n = \overline{1, N}$, and the average length of the information message $\tau_n \in \square$, $n = \overline{1, N}$.

If the communication resource allocation of volume V cannot be evenly distributed among requests while adhering to the guaranteed resource volume v, then the incoming request is redirected to a buffer with a capacity of w. Within this concept, we denote the maximum number of requests that the base station can simultaneously handle as $\lfloor V/v \rfloor = K$. Requests redirected to the buffer may wait for service for an extended period, thereby leaving the system at an intensity $\mu_n > 0$, individually denoted $\forall n \in \mathrm{N}$.

Let's characterize the above-defined process using stochastic dependence $K(t) \in \{0, \ldots, \lfloor V/v \rfloor\}$, representing the number of requests supported by the base station at the moment $t \geq 0$. The state space determined by the dependence $K(t)$ is identified as $Y := \{k \in \{0, \ldots, K, \ldots, \min(N, K + w)\}\}$. Depending on the relationship between the available communication resource volume at the base station and the number of active subscribers, the investigated process can evolve according to three scenarios:

1. The number of incoming requests from active subscribers is less than the number of requests the base station can handle using the available communication resource volume: $0 < N \leq K$. Therefore, all incoming requests will be accepted for service.
2. The number of incoming requests from active subscribers is greater than the number of requests the base station can handle using the available communication resource volume: $K < N \leq K + w$. Therefore, incoming requests are redirected to the buffer.
3. The number of incoming requests from active subscribers is greater than both the number of requests the base station can handle using the available communication resource volume and the buffer capacity: $N > K + w$. Therefore, incoming requests will be lost.

The system of equilibrium equations for the investigated process, represented by the dependence $K(t)$, taking into account the scenarios of its development described above, is given by the form

$$
\begin{cases}
N\eta q_0 = Vq_1/\tau, \\
\left((N-k)\eta + V/\tau\right)q_k = (N-k+1)\eta q_{k-1} + \\
+Vq_{k+1}/\tau \, \forall k = \overline{1, (K-1)}, \\
\left((N-k)\eta + V/\tau + (k-K)\mu\right)q_k = (N-k+1)\eta q_{n-1} + \\
+\left(V/\tau + (k+1-K)\mu\right)q_{k+1} \forall k = \overline{K, (K+w-1)}, \\
\left(V/\tau + w\mu\right)q_{K+w} = (N-K-w+1)\eta q_{K+w-1},
\end{cases}
\tag{3}
$$

where q_k denotes the sought stationary probabilities of the investigated process being in the k-th states, $k \in \{0, \ldots, K, \ldots, \min(N, K+w)\}$.

From the system of equations (3), we express the stationary probability distribution as

$$q_k = \begin{cases} (\eta\tau/V)^k \, p_N^k q_0 \forall k = \overline{1, \min(K,N)}, \\ \dfrac{(\tau/V)^K \eta^k}{\prod_{i=1}^{k-K}\left(\dfrac{V}{\tau} + \mu i\right)} p_N^k q_0 \forall k = \overline{(K+1), \min(K+w, N)}, \end{cases} \tag{4}$$

where

$$q_0 = 1 \Bigg/ \left(\sum_{k=0}^{\min(K,N)} \left(\frac{\eta\tau}{V}\right)^k p_N^k + \left(\frac{\tau}{V}\right)^K \sum_{k=1}^{\min(w,N-K)} \frac{\eta^k p_N^k}{(V/\tau) + \mu i} \right)$$

is involved.

At known values of the indicators in (4) for the target service-oriented mMTC traffic handling system in a 5G cluster, it is possible to calculate metrics of performance indicators, including the probability Q of losing an incoming request due to system overload, the average duration T_{sys} of a request's stay in the system (including the average service duration T_{serv} and the average duration of a request's stay in the buffer T_{buf}), and the average number of requests in the system R_{sys} (including the average number of requests in the service stage R_{serv} and the average number of requests in the buffer R_{buf}):

$$Q = \begin{cases} 0 \forall 0 < N \le K + w, \\ q_{K+w} \forall N > K + w, \end{cases} \tag{5}$$

$$T_{sys} = \begin{cases} \left(\sum_{i=0}^{N} i q_i\right) \Bigg/ \left(\sum_{k=0}^{N-1} (N-k)\eta q_k\right) \forall 0 < N \le K + w, \\ \left(\sum_{i=0}^{K+w} i q_i\right) \Bigg/ \left(\sum_{k=0}^{K+w-1} (N-k)\eta q_k\right) \forall N > K + w, \end{cases} \tag{6}$$

$$T_{serv} = \begin{cases} \dfrac{\sum_{i=0}^{N} i q_i}{\sum_{k=0}^{N-1} (N-k)\eta q_k} \forall 0 < N \le K, \\ \dfrac{\sum_{i=0}^{K} i q_i + K \sum_{i=1}^{N-K} q_{K+i}}{\sum_{k=0}^{N-1} (N-k)\eta q_k} \forall K < N \le K + w, \\ \dfrac{\sum_{i=0}^{K} i q_i + K \sum_{i=1}^{w} q_{K+i}}{\sum_{k=0}^{K+w-1} (N-k)\eta q_k} \forall N > K + w, \end{cases} \tag{7}$$

$$T_{buf} = \begin{cases} \left(\sum_{i=1}^{N-K} i q_{K+i}\right) \Bigg/ \left(\sum_{k=0}^{N-1} (N-k)\eta q_k\right) \forall K < N \le K + w, \\ \left(\sum_{i=1}^{w} i q_{K+i}\right) \Bigg/ \left(\sum_{k=0}^{K+w-1} (N-k)\eta q_k\right) \forall N > K + w, \end{cases} \tag{8}$$

$$R_{sys} = \begin{cases} \sum_{i=0}^{N} iq_i \forall 0 < N \le K + w, \\ \sum_{i=0}^{K+w} iq_i \forall N > K + w, \end{cases} \quad (9)$$

$$R_{serv} = \begin{cases} \sum_{i=0}^{N} iq_i \forall 0 < N \le K, \\ \sum_{i=0}^{K} iq_i + K \sum_{i=1}^{N-K} q_{K+i} \forall K < N \le K + w, \\ \sum_{i=0}^{K} iq_i + K \sum_{i=1}^{w} q_{K+i} \forall N > K + w, \end{cases} \quad (10)$$

$$R_{buf} = \begin{cases} \sum_{i=1}^{N-K} iq_{K+i} \forall K < N \le K + w, \\ \sum_{i=1}^{w} iq_{K+i} \forall N > K + w. \end{cases} \quad (11)$$

Results and Discussion

We will apply the capabilities of the simulation modelling method to evaluate the typical instance of the service-oriented mMTC traffic handling system in a 5G cluster in the qualitative metrics (5)-(11).

To calculate the indicators (4) by solving the system of equilibrium equations (3), we will assign values to a series of input parameters specific to the investigated instance of the 5G cluster, including: - The number of subscriber terminals $N = 80$; - Guaranteed volume of communication resources allocated by the base station to serve an accepted request $v = 512$ Kbps; - Parameters of the intensity of the exponentially distributed input request flow $\eta = [0.01;30]$; - The mathematical expectation of the exponentially distributed stochastic length of an information message $\tau = 2048$ Kb; - The intensity of requests leaving the buffer due to excessive waiting time for service $\mu = 10^{-7}$; - The buffer capacity (in the number of requests) $w = 30$.

Based on the defined input parameters, we solve the system of equilibrium equations (3) using the Gaussian method to determine the probabilities (4). As a result, all the values necessary for calculating the qualitative metric indicators (5)-(11) are determined.

To demonstrate the informativeness of the proposed qualitative metric using expressions (9)-(11), we will calculate the dependencies $\{R_{sys}, R_{serv}, R_{buf}\} = f(\eta)$ on the specified total volume of $V = 16$ Mbps of communication resources of the base station for the investigated instance of the 5G cluster. The calculation results, presented in the form of graphs, are visualized in Figure 1.

The graphs presented in Fig. 1 show that the quality indicators (9)-(11) are sensitive to the increasing intensity of incoming requests from subscriber terminals to the base station of the investigated 5G cluster. Under the given initial parameters, buffer overflow is observed at $\eta \ge 0.6$, synchronously reflected in the stabilization of the number of serviced requests in the system (the graph of dependency $R_{sys} = f(\eta)$).

An increase in the intensity η beyond the value of 0.6 is accompanied by the loss of new incoming requests, as the system buffer is filled, and all available communication resources are allocated to serve the accepted requests. It is worth noting separately that the nature of the graphs presented in Fig. 1 and beyond corresponds to and is largely determined by the distribution of the stochastic variable-argument.

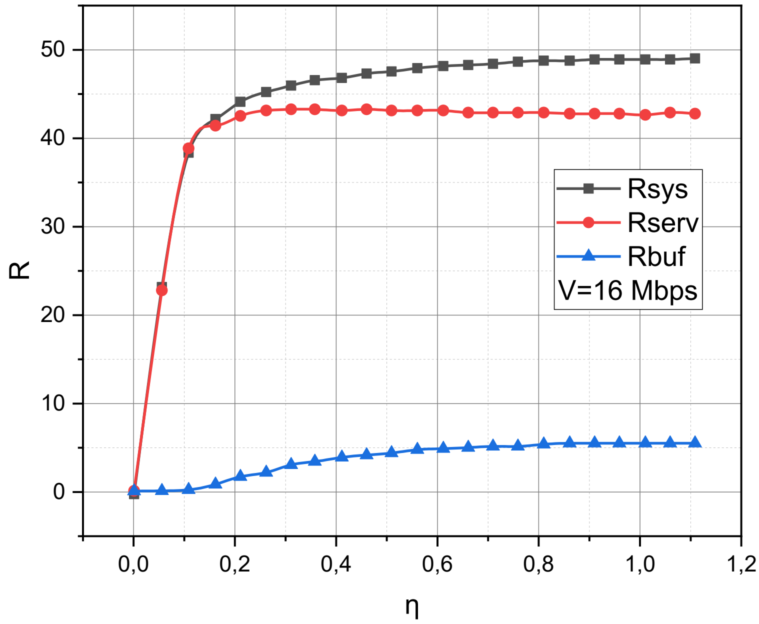

Figure 1: The graphs of dependency $\{R_{sys}, R_{serv}, R_{buf}\} = f(\eta)$, calculated at $V = 16$ Mbps.

We complement the information presented in Fig. 1 by calculating the quality indicators (6)-(8) for the investigated instance of the 5G cluster. The calculation results, visualized in the form of the graphs of dependency $\{T_{sys}, T_{serv}, T_{buf}\} = f(\eta)$ at $V = 16$ Mbps, are presented in Fig. 2.

As expected, the shape and dynamics of the graphs of dependency $\{T_{sys}, T_{serv}, T_{buf}\} = f(\eta)$ presented in Fig. 2 are similar to the results shown in Fig. 1.

We observe that after $\eta = 0.6$, the operation of the investigated 5G cluster stabilizes for all its components (subscriber terminals (request generators), a base station (servicing device for accepted requests), buffer (means of reducing the number of lost requests)).

However, it should be noted that the plateau saturation of all graphs in Fig. 2 at $\eta \geq 0.6$ indicates that the assigned values of the output parameters μ and w are too small, preventing the base station from processing the parameterized flow of incoming requests without losses.

This issue can be addressed by increasing either the value of μ (the characteristic parameter of subscriber terminals), or the value of w (characteristic parameter of the base station), or the values of both mentioned parameters.

From Fig. 1 and Fig. 2, it is evident that the specified value of $V = 16$ Mbps is insufficient for the effective operation of the 5G cluster, as defined by the corresponding configuration of output parameter values. Let's explore how the behaviour of the investigated communication system will change if the value of the parameter V significantly increases: $V = 2.0$ Gbps.

We will conduct a study similar to the one depicted in Fig. 1 and Fig. 2.

The calculation results of the dependencies $\{R_{sys}, R_{serv}, R_{buf}\} = f(\eta)$ and $\{T_{sys}, T_{serv}, T_{buf}\} = f(\eta)$ at $V = 2.0$ Gbps are presented in Fig. 3 and Fig. 4, respectively.

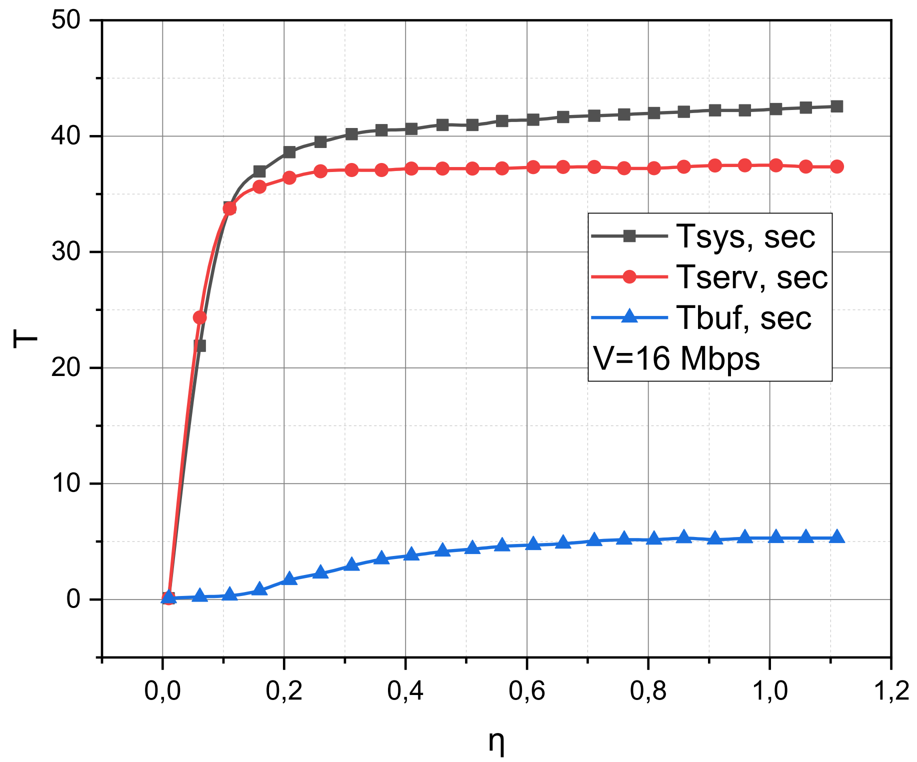

Figure 2: The graphs of dependency $\{T_{sys}, T_{serv}, T_{buf}\} = f(\eta)$, calculated at $V = 16$ Mbps.

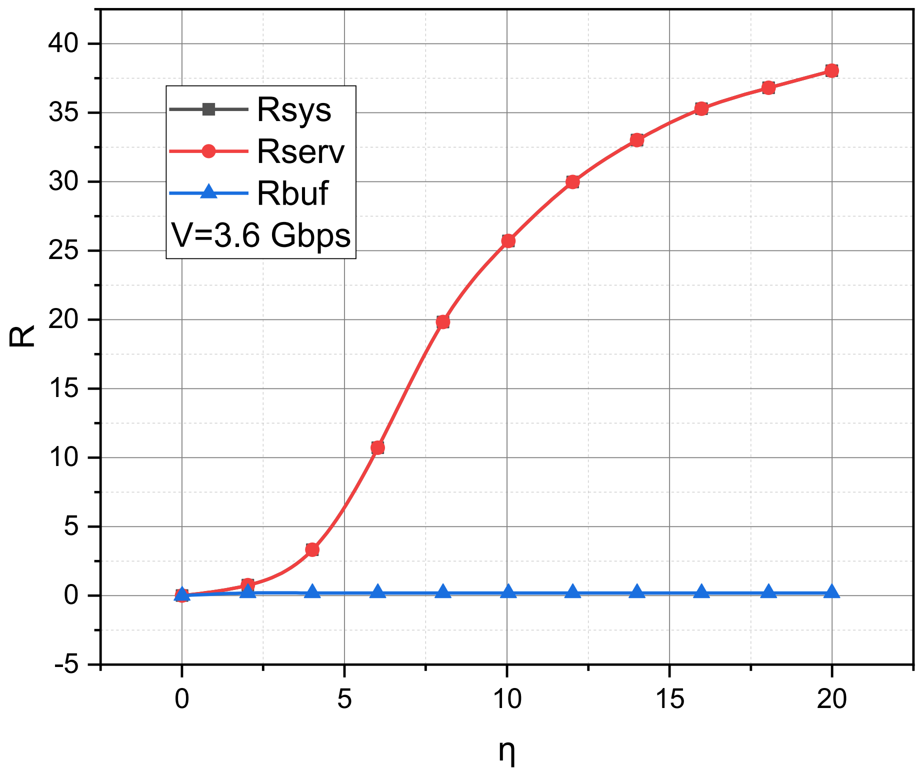

Figure 3: The graphs of dependency $\{R_{sys}, R_{serv}, R_{buf}\} = f(\eta)$, calculated at $V = 2.0$ Gbps.

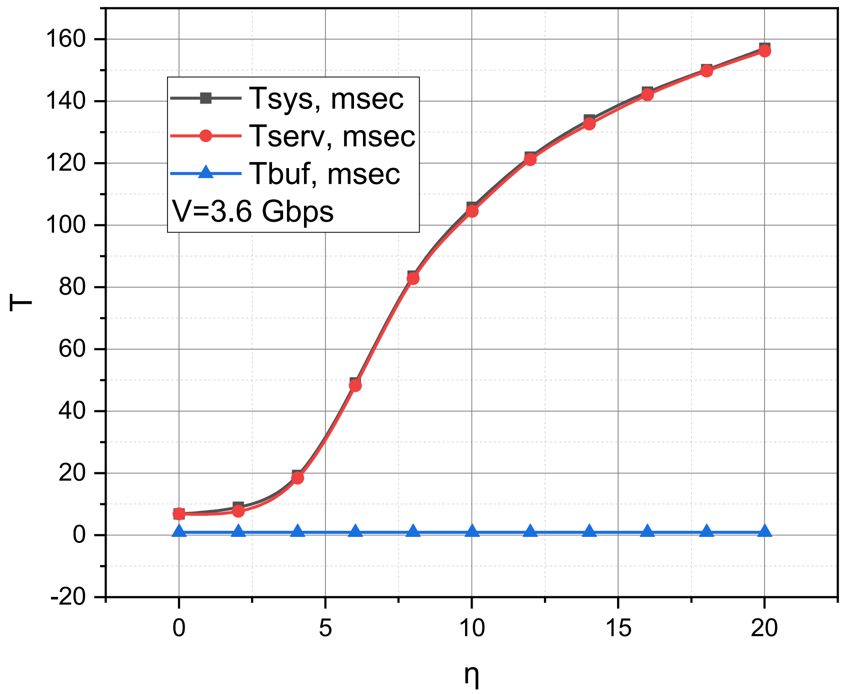

Figure 4: The graphs of dependency $\{T_{sys}, T_{serv}, T_{buf}\} = f(\eta)$, calculated at $V = 2.0$ Gbps.

Let's analyze the results presented in Fig. 3 and Fig. 4. In contrast to the uniform graphs shown in Fig. 1 and Fig. 2, the graphs in Fig. 3 and Fig. 4 have an approximately linear character. This indicates that the specified value of $V = 2.0$ Gbps is sufficient for servicing the flow of incoming requests from $V = 3.6$ subscriber terminals directed towards the base station. The flow of incoming requests is serviced without losses across the entire investigated range of argument values η.

In addition to the character of the graphs of dependencies $\{R_{sys}, R_{serv}\} = f(\eta)$, $\{T_{sys}, T_{serv}\} = f(\eta)$, this is evidenced by the fact that the buffer is not utilized (the $R_{buf} = f(\eta)$, $T_{buf} = f(\eta)$ graphs are parallel to the x-axis with ordinate values near zero).

The obtained results indicate that for the investigated range of intensity values η, the specified value of $V = 2.0$ Gbps is excessive for the 5G cluster under consideration. This brings us back to the point that the value of the parameter V for the target information and communication system should be chosen not empirically but as a result of solving the optimization problem with the objective function (1), as outlined in Section 2.1.

Conclusions

In the list of technologies proposed by the 5G/IMT-2020 standards specification [1], the support for massive Machine Type Communication (mMTC) holds a prominent position, and this is not coincidental. The implementation of Industry 4.0 leads to a continuous expansion of sensor network coverage areas, connecting them to the cloud via wired communication is neither technologically nor economically feasible, and often simply impossible. The mMTC technology is specifically designed to address a plethora of issues related to organizing informational interactions around sensor networks. However, any system operates predictably and efficiently only if its structure is transparent and optimal in the context of its intended purpose. Therefore, the research task of the rational organization of the data collection process from multiple sensor networks with a finite number of terminals, located in the vicinity of a 5G cluster, is relevant and of great practical significance.

To address the research task, the article proposes a Markov model of the controlled process of uplink transfer of segmented informational messages from a finite set of terminals belonging to different sensor networks deployed in the coverage area of the target 5G cluster. The segregation of sensor networks at the model level is ensured by their service orientation in dedicated virtual network segments organized in the information environment of the 5G base station. Moreover, the model allows for establishing priorities for services characteristic of each sensor network. The specificity of mMTC traffic is taken into account by the base station, which accepts an incoming request only if there is a minimum guaranteed amount of available communication resources necessary for its processing. It is assumed that an unaccepted request may wait for processing in a buffer of finite capacity, which it leaves either upon acceptance for service or upon expiration of the assigned waiting time.

The model serves as the basis for formalizing the metrics of quality indicators, including the probability of losing an incoming request due to system overload, indicators of the average duration of a request's stay in the system, and indicators of the average number of requests in the system. To demonstrate the functionality of the proposed mathematical framework for the target service-oriented traffic handling system of mMTC subscribers in a 5G cluster, dependencies of quality metric indicators on the intensity of incoming requests to the base station were calculated. The obtained results allowed for the estimation of the volume of communication resources that can be reasonably predicted for the efficient operation of the target 5G cluster with mMTC traffic.

Further research is planned to focus on finding variations in the formulations of optimization problems (1), (2) capable of reflecting the specifics of different operating modes of the instance of the service-oriented traffic handling system for mMTC subscribers in a 5G cluster. Additionally, the intention is to compare and extend the results presented in [21, 22].

Acknowledgements

The authors are grateful to all colleagues and institutions that contributed to the research and made it possible to publish its results.

Funding

This research is part of the project No. 2022/45/P/ST7/03450 co-funded by the National Science Centre and the European Union Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie grant agreement No. 945339.

References

[1] M. Fuentes et al., “5G New Radio Evaluation Against IMT-2020 Key Performance Indicators,” IEEE Access, vol. 8. Institute of Electrical and Electronics Engineers (IEEE), pp. 110880–110896, 2020. doi: 10.1109/access.2020.3001641.

[2] V. Tyagi and S. Singh, “Network resource management mechanisms in SDN enabled WSNs: A comprehensive review,” Computer Science Review, vol. 49. Elsevier BV, p. 100569, Aug. 2023. doi: 10.1016/j.cosrev.2023.100569.

[3] V. S. Pana, O. P. Babalola, and V. Balyan, “5G radio access networks: A survey,” Array, vol. 14. Elsevier BV, p. 100170, Jul. 2022. doi: 10.1016/j.array.2022.100170.

[4] N. Sharma and K. Kumar, “Resource allocation trends for ultra dense networks in 5G and beyond networks: A classification and comprehensive survey,” Physical Communication, vol. 48. Elsevier BV, p. 101415, Oct. 2021. doi: 10.1016/j.phycom.2021.101415.

[5] L. Miuccio et al., “A QoS-aware and channel-aware Radio Resource Management framework for multi-numerology systems,” Computer Communications, vol. 191. Elsevier BV, pp. 299–314, Jul. 2022. doi: 10.1016/j.comcom.2022.05.009.

[6] R. Kumar, D. Sinwar, and V. Singh, “QoS aware resource allocation for coexistence mechanisms between eMBB and URLLC: Issues, challenges, and future directions in 5G,” Computer Communications, vol. 213. Elsevier BV, pp. 208–235, Jan. 2024. doi: 10.1016/j.comcom.2023.10.024.

[7] X. Tong, X. Li, and Y. Liu, “An auction-based distributed network slicing scheme for resource allocation in satellite-UAV integrated networks,” Computer Communications, vol. 210. Elsevier BV, pp. 58–68, Oct. 2023. doi: 10.1016/j.comcom.2023.08.002.

[8] J. Huang et al., “Opportunistic capacity based resource allocation for 6G wireless systems with network slicing,” Future Generation Computer Systems, vol. 140. Elsevier BV, pp. 390–401, Mar. 2023. doi: 10.1016/j.future.2022.10.032.

[9] L. Zhang, B. Cao, and G. Feng, “Opportunistic admission and resource allocation for slicing enhanced IoT networks,” Digital Communications and Networks, vol. 9, no. 6. Elsevier BV, pp. 1465–1476, Dec. 2023. doi: 10.1016/j.dcan.2022.08.007.

[10] K. Obelovska, O. Panova, and V. Karovič Jr., “Performance Analysis of Wireless Local Area Network for a High-/Low-Priority Traffic Ratio at Different Numbers of Access Categories,” Symmetry, vol. 13, no. 4. MDPI AG, p. 693, Apr. 15, 2021. doi: 10.3390/sym13040693..

[11] J. S. Camargo et al., “Dynamic slicing reconfiguration for virtualized 5G networks using ML forecasting of computing capacity,” Computer Networks, vol. 236. Elsevier BV, p. 110001, Nov. 2023. doi: 10.1016/j.comnet.2023.110001.

[12] R. Shone, K. Glazebrook, and K. G. Zografos, “Resource allocation in congested queueing systems with time-varying demand: An application to airport operations,” European Journal of Operational Research, vol. 276, no. 2. Elsevier BV, pp. 566–581, Jul. 2019. doi: 10.1016/j.ejor.2019.01.024.

[13] J. Niño-Mora, “Resource allocation and routing in parallel multi-server queues with abandonments for cloud profit maximization,” Computers & Operations Research, vol. 103. Elsevier BV, pp. 221–236, Mar. 2019. doi: 10.1016/j.cor.2018.11.012.

[14] N. Dahlquist, A. Saradagi, and G. Nikolakopoulos, “Reactive Task Allocation for Balanced Servicing of Multiple Task Queues,” IFAC-PapersOnLine, vol. 56, no. 2. Elsevier BV, pp. 9771–9776, 2023. doi: 10.1016/j.ifacol.2023.10.293.

[15] U. Singh et al., “Scalable priority-based resource allocation scheme for M2M communication in LTE/LTE-A network,” Computers and Electrical Engineering, vol. 103. Elsevier BV, p. 108321, Oct. 2022. doi: 10.1016/j.compeleceng.2022.108321.

[16] A. Mishra, A. Swain, A. K. Ray, and R. M. Shubair, “HetNet/M2M/D2D communication in 5G technologies,” 5G IoT and Edge Computing for Smart Healthcare. Elsevier, pp. 45–87, 2022. doi: 10.1016/b978-0-323-90548-0.00002-4.

[17] U. Singh et al., “Scalable priority-based resource allocation scheme for M2M communication in LTE/LTE-A network,” Computers and Electrical Engineering, vol. 103. Elsevier BV, p. 108321, Oct. 2022. doi: 10.1016/j.compeleceng.2022.108321.

[18] K. Boutiba, M. Bagaa, and A. Ksentini, “Optimal radio resource management in 5G NR featuring network slicing,” Computer Networks, vol. 234. Elsevier BV, p. 109937, Oct. 2023. doi: 10.1016/j.comnet.2023.109937.

[19] R. Kumar, D. Sinwar, and V. Singh, “QoS aware resource allocation for coexistence mechanisms between eMBB and URLLC: Issues, challenges, and future directions in 5G,” Computer Communications, vol. 213. Elsevier BV, pp. 208–235, Jan. 2024. doi: 10.1016/j.comcom.2023.10.024.

[20] K. F. Muteba, K. Djouani, and T. Olwal, “5G NB-IoT: Design, Considerations, Solutions and Challenges,” Procedia Computer Science, vol. 198. Elsevier BV, pp. 86–93, 2022. doi: 10.1016/j.procs.2021.12.214.

[21] V. Kovtun and K. Grochla, “Investigation of the competitive nature of eMBB and mMTC 5G services in conditions of limited communication resource,” Scientific Reports, vol. 12, no. 1. Springer Science and Business Media LLC, Sep. 26, 2022. doi: 10.1038/s41598-022-20135-5.

[22] V. Kovtun, I. Izonin, and M. Gregus, “Modeling a session of subject-system interaction in a wireless communication infrastructure with a mixed resource,” PLOS ONE, vol. 17, no. 7. Public Library of Science (PLoS), p. e0271536, Jul. 18, 2022. doi: 10.1371/journal.pone.0271536.