electronics

Article

Investigation of the Information Interaction of the Sensor Network End IoT Device and the Hub at the Transport Protocol Level

Viacheslav Kovtun *, Krzysztof Grochla and Konrad Połys

Internet of Things Group, Institute of Theoretical and Applied Informatics Polish Academy of Sciences, Bałtycka 5, 44-100 Gliwice, Poland; kgrochla@iitis.pl (K.G.); kpolys@iitis.pl (K.P.)
* Correspondence: kovtun_v_v@vntu.edu.ua

Abstract: The study examines the process of information transfer between the sensor network end IoT device and the hub at the transport protocol level focused on using the 5G platform. The authors interpreted the researched process as a semi-Markov (focused on the dynamics of the size of the protocol sliding window) process with two nested Markov chains (the first characterizes the current size of the sliding window, and the second, the number of data blocks sent at the current value of this characteristic). As a result, a stationary distribution of the size of the sliding window was obtained both for the resulting semi-Markov process and for nested Markov chains, etc. A recursive approach to the calculation of the mentioned stationary distribution is formalized. This approach is characterized by linear computational complexity. Based on the obtained stationary distribution of the size of the sliding window, a distribution function is formulated that characterizes the bandwidth of the communication channel between the entities specified in the research object. Using the resulting mathematical apparatus, the Window Scale parameter of the TCP Westwood+ protocol was tuned. Testing has shown the superiority of the modified protocol over the basic versions of the BIC TCP, TCP Vegas, TCP NewReno, and TCP Veno protocols in conditions of data transfer between two points in the wireless sensor network environment.

Keywords: information and communication technologies; data transfer; transport protocol; end IoT device; hub; sliding window size; bandwidth

Citation: Kovtun, V.; Grochla, K.; Połys, K. Investigation of the Information Interaction of the Sensor Network End IoT Device and the Hub at the Transport Protocol Level. *Electronics* **2023**, *12*, 4662. https://doi.org/10.3390/electronics12224662

Academic Editors: Dionisis Kandris and Eleftherios Anastasiadis

Received: 31 October 2023
Revised: 9 November 2023
Accepted: 13 November 2023
Published: 15 November 2023

1. Introduction

A key role in managing modern network traffic belongs to transport layer protocols (in particular the TCP (Transmission Control Protocol) [1–4]). By controlling network connections at a point-to-point level, the main algorithms of these protocols form both quantitative and qualitative characteristics of bidirectional packet flow following the physical characteristics and level of congestion of a network route used. Therefore, it is a property of transport protocols that make a decisive contribution to ensuring the reliability, stability and performance of data networks. The latter makes the task of modeling the behavior and analyzing the performance of transport protocols especially the TCP protocol very relevant. We separately, note that the study of the properties of the TCP protocol in various application scenarios is relevant, including because more than 95% of all data flows in the world are controlled by this protocol [5–7].

The Internet of Things (IoT) is built on existing network infrastructure, technologies and protocols currently used in homes, offices and the Internet. This means that most IoT runs on existing TCP/IP networks. TCP/IP uses a four-layer model with specific protocols at each layer. The diagram below (Figure 1) shows a comparison of the protocols currently in use and those most likely to be used for IoT [8–10] in the future.

Modern Internet protocols	Current and expected IoT protocols
Application layer	
HTTP (FTP, SMTP, IMAP)	MQTT (COAP, AMQP)
Transport layer	
TCP, UDP	TCP, UDP
Network layer	
IPv4, IPv6	IPv6, IPv4
Communication layer	
Ethernet, Wi-Fi, GSM	Ethernet, Wi-Fi, GSM, LTE-M, Lora, SigFox
Protocol	
TCP/IP model Internet and IoT protocols	

Figure 1. Analysis of the most current protocols used in the TCP/IP model to support Internet and IoT data flows.

Figure 1 shows that most changes will occur at the communication and application layers, while the network and transport layers are likely to remain unchanged. Thus, by solving the problems of optimizing the TCP protocol for the specific needs of IoT, we are engaged in promising research, the results of which will be in demand not only in the present but also in the future.

Note that TCP creates end-to-end connections on top of the inherently unreliable and best-effort IP packet service through a unique procedure known as a "three-way handshake". During this process, the client sends three TCP segments to the server, and the server responds, establishing a connection. However, due to the nature of IP routing networks, packets containing the TCP segment requesting a new connection and the server's response can sometimes get lost, leading to uncertainty for the communicating hosts. The third message in the sequence enhances the overall reliability of the connection. TCP employs distinct terminology for its connection establishment process. It utilizes a solitary bit known as the SYN (SYNchronization) bit to signify a connection request. This single bit is encapsulated within a comprehensive 20-byte (typically) TCP header, and additional data, including the Initial Sequence Number (ISN) for segment tracking, is transmitted to the receiving host. ACKnowledgments for connections and data segments are confirmed using the ACK bit, while a request to conclude a connection is conveyed through the FIN (FINal) bit. It's important to highlight that when transmitting a single request and response pair within segments, TCP necessitates the creation of an additional seven packets. This results in a substantial packet overhead, and the entire procedure tends to be sluggish when operating over high-latency (delayed) connections. This is a contributing factor to the growing popularity of UDP, especially as networks continue to improve their reliability.

However, let us focus on the network and transport layers in more detail. At the network layer, IPv6 will dominate in the long term. It is unlikely that IPv4 will be used, but it may play a role in the initial stages. Most IoT devices for the home, such as smart light bulbs, currently use IPv4. TCP dominates the transport layer on the Internet. It is used in both HTTP and many other popular Internet protocols (SMTP, POP3, IMAP4, etc.). MQTT (Message Queuing Telemetry Transport) protocol is expected to become one of the main application layer protocols for messaging. Currently, MQTT uses TCP. However, there is an opinion [8] that in the future, due to lower overhead costs, the share of using UDP to serve IoT needs will grow (MQTT-SN operating on top of UDP will probably become more widespread). At the same time, the share of IoT devices that simultaneously operate both on the Internet and on IoT is growing. This statement is confirmed by data on protocol support for IoT platforms: Microsoft Azure (MQTT, AMQP, HTTP and HTTPS), AWS (MQTT, HTTPS, MQTT over WebSockets), IBM Bluemix (MQTT, HTTPS, MQTT), Thingworx (MQTT, HTTPS, MQTT, AMQP). Thus, the trinity of the Internet, IoT, and

TCP will be stable and inseparable in the future. At the same time, we will mention the main disadvantages of TCP, namely, the difficulty in setting up and managing specific use cases; the protocol does not guarantee the delivery of data packets. It is the optimization of TCP protocol parameters to eliminate these shortcomings in the context of the interaction between the "sensor network end IoT device" and the "hub" that is the motivation for the research presented below.

2. State-of-the-Art

The chosen subject area is characterized by a high intensity of research. As confirmation of this fact, we mention both highly cited [1–9] and new research works [10–15].

In [10], T. Toprasert and W. Lilakiataskun introduce a Markov Decision Process (MDP) aimed at improving congestion avoidance. The researchers argue that the MIMD mechanism surpasses both TCP-Illinois and TCP-Scalable in managing window size congestion. Drawing from their findings, they presented a new iteration of the TCP protocol named TCP-Siam. This protocol incorporates a coefficient designed to optimize the congestion window (cWnd), enhancing performance when packets are lost over loss-prone links in a WMN.

In [11], Hurni and colleagues investigated methods to enhance TCP performance. They delved into the impacts of distributed caching coupled with local retransmission techniques. In this approach, every intermediary node stores TCP segments and retransmits the segment if its ACK (acknowledgement) is notably delayed, determined by RTT (round-trip time) estimates. They incorporated their solution into the Contiki OS's uIP stack with a module they named "caching and congestion control" (cctrl). This was then tested across several radio-duty cycling MAC (medium access control) protocols using a real-world testbed of seven TelosB motes. While experiments revealed the cctrl module boosted TCP throughput in numerous settings, its efficacy largely hinged on the specific RDC MAC protocols employed.

Kim et al., as mentioned in [12], undertook an experimental analysis of TCP performance over RPL within an IPv6-driven testbed. This testbed, a low-power and lossy network (LLN), incorporated 30 TelosB devices operating on the TinyOS BLIP stack, complemented by one LBR (LoWPAN border router) and a Linux server. Their observations revealed that TCP displayed notable throughput disparities across nodes in multi-hop LLNs. Moreover, RPL's lack of consideration for traffic load balancing could potentially deteriorate TCP performance.

In an effort to rectify the TCP fairness issues among LLN endpoints, Park and Paek introduced TAiM (TCP assistant in the middle) in [13]. This system intervenes mid-way in TCP communication, specifically at the LBR, and adjusts the RTT of ongoing flows. TAiM holds onto packets and deliberately introduces a delay before forwarding them. This means a flow with reduced throughput experiences a briefer delay, while a flow with increased throughput faces a more extended delay. Experiments utilizing the BLIP stack indicated that TAiM enhanced TCP fairness without compromising the overall throughput.

Gomez et al., in [14], offered guidelines for streamlined TCP implementation, optimized for IoT contexts. Specifically, when considering the RTO algorithm, they advocated for the adoption of CoCoA within TCP.

With the advent of advanced low-power embedded devices boasting greater processing capabilities and increased memory, Kumar et al. demonstrated that a comprehensive TCP can comfortably operate within the CPU and memory limits of contemporary wireless sensor network platforms. They achieved this by implementing a full-scale TCP named TCPlp [15], which draws from the complete capabilities of the TCP in the FreeBSD OS.

As the main results in most of the mentioned works, there are various estimates of the stationary mathematical expectation of the bandwidth of the communication channel, which is controlled by the transport protocol (directly or consolidated). There is also a study of the second moment of bandwidth [12–15] for a point stochastic process, specially selected within the original process of information transfer. Such a trend is quite understandable because, among the typical tasks that are solved in the field of information and communication technologies, there is not only the determination of the state of the communication process at an arbitrary moment in time but also the forecasting of its development and the planning of resources “for growth”. It is in such problems that the highest points of the distributions of controlled parameters, as well as their quantiles, are of greatest interest.

Most of the mentioned estimates of the characteristic parameters of the information interaction process are formalized in the form of simple algebraic constructs, in which the efficiency and simplicity of calculations are balanced by the generality of the approximation of the obtained values. In particular, general assessments completely ignore several significant features that appear in situations where the process of information exchange in sensor networks is observed. Let us formulate the most obvious of these features:

- service signals regarding the result of the transfer of a packet of data blocks arrive at random moments, which leads to the interpretation of RTT as a stochastic parameter with a known distribution, depending on the size of the sliding window of the transport protocol. This feature represents the basic specificity of the Ultra-Reliable Low Latency Communication (URLLC) technology as a component of the 5G platform;
- when forming the parametric space of the model of the information exchange process, one should take into account the fact that in real information and communication systems, both the size of the sliding window and the bandwidth of the communication channel are large but always finite. This feature represents the basic specificity of the Massive Machine-Type Communications (mMTC) technology as a component of the 5G platform;
- both the distribution of the size of the sliding window and the distribution of bandwidth should be defined in terms of the ratio of RTT and bandwidth of the communication channel. This approach will make it possible to eliminate the potential influence of the speed characteristics of the data transfer channel on the adequacy of the description of the studied process by the created mathematical apparatus. This feature represents the basic specificity of the enhanced Mobile BroadBand (eMBB) technology as a component of the 5G platform.

Taking into account the strengths and weaknesses of the mentioned methods, we will formulate the necessary attributes of scientific research.

Studied object: The object of our research is the transport layer for managing the process of data transfer between the sensor network end IoT device and the hub using the communication capabilities of the 5G platform.

Research subject: the probability theory and mathematical statistics, the stochastic processes theory, the queuing theory, and the experiment planning theory.

The aim of the research: is to formalize the process of information transfer between the sensor network end IoT device and the hub at the transport level with the determination of the essential characteristic parameters of the protocol.

Research objectives:

- to formalize the process of information transfer between the sensor network end IoT device and the hub based on the stochastic processes theory and the queuing theory;
- to formalize in the analytical basis of the researched process the stationary distribution of the size of the sliding window as a characteristic parameter that determines the intensity of the information flow from the addressee;
- to formalize in the analytical basis of the researched process the distribution function of the bandwidth of the communication channel between the entities specified in the research object;
- justify the adequacy of the proposed mathematical apparatus and demonstrate its functionality with an example.

Main contribution. The study examines the process of information transfer between the sensor network end IoT device and the hub at the transport protocol level. In this context, a queuing system with controlled input flow, deterministic service, feedback and an unlimited queue is synthesized. The authors interpreted the research object as a semi-Markov (focused on the dynamics of the size of the protocol sliding window) process with two nested Markov chains (the first characterizes the current size of the sliding window, and the second—the number of data blocks sent at the current value of this characteristic). As a result, a stationary distribution of the size of the sliding window (a parameter that determines the intensity of the information flow from the addressee) was obtained both for the resulting semi-Markov process and for nested Markov chains, etc. A recursive approach to the calculation of the mentioned stationary distribution is formalized. This approach is characterized by linear computational complexity. Based on the obtained stationary distribution of the size of the sliding window, a distribution function is formulated that characterizes the bandwidth of the communication channel between the entities specified in the research object.

One of the key algorithms in the family of TCP-like transport protocols is an algorithm conventionally called Additive-Increase/Multiplicative-Decrease (AIMD) [16–18]. The AIMD algorithm is focused on increasing the intensity of the flow of packets generated by the sender if the recipient confirms their successful delivery. The mathematical model presented in Section 3 reflects the impact of the AIMD algorithm on massive traffic between the sensor network end IoT device and the hub. Also in Section 3, an analytical form of bandwidth distribution for such a connection is obtained. Section 4 presents the results of the experiments, describing the equipment used and the technologies used to register the empirical data. This section presents the results of comparing the TCP Westwood+ protocol, the Window Scale parameter of which was determined based on the author's mathematical apparatus, with BIC TCP, TCP Vegas, TCP NewReno, TCP Veno without tuning. In Section 5, conclusions are drawn taking into account the results obtained, and directions for further research are formulated.

3. Materials and Methods

Regardless of the type of operating system, the modern transport protocol is designed to ensure the reliable reception of data blocks sent by the communication channel by the addressee. The TCP protocol uses the sliding window to regulate how many packets are in transit to maximize the transmission throughput assuring the reliability of the communication. A few different algorithms have been proposed to regulate the window size, starting from TCP Reno and NewReno to the TCP BIC and CUBIC used in modern operating systems. Suppose that the size of the sliding window is equal to $l > 0$. The basic mechanism of the regulation of the window size can be modeled as follows: if the sender received confirmation from the addressee about the successful receipt of the data block packet at the current sliding window size, then the sliding window size for the next in line to send the data block packet will be increased to the value $l + \lfloor l/n \rfloor$, where $n \geq 2$, $n \in \mathbb{N}$. Otherwise, the size of the sliding window will be reduced to the value $\lfloor l/n \rfloor$. While this is a simplified model mimicking the behavior of the traditional TCP Reno algorithm, the more advanced window size control algorithms, such as BIC and CUBIC can still be approximated with it.

Let us generalize the probabilities f_i that $i = 1, 2, \ldots$ consecutively sent data blocks will be received by the addressee in the form of a distribution of $\{f_i\}$. We consider the stochastic elements of this set to be independent. Compliance with this condition allows us to classify the entity $\{f_i\}$ as a geometric distribution, the parameter $p \in (0, 1)$ of which characterizes the probability of losing a data block during the transfer process.

We will assume that the circular delay D is constant and equal to one. Under the condition of guaranteed successful transfer of all data blocks sent by n RTT, the protocol will send

$$N(n) = 1 + 2 + \ldots + n = \frac{1}{2}\left(n^2 + n\right) \quad (1)$$

data blocks starting from the size of the sliding window $l = 1$. Therefore, the amount of data equal to $N(n)$ will be sent in $N(t) = O\left(t^2\right)$ units of time. The non-linear growing character of the function $N(t)$ prompts the introduction of a parameter whose value will limit this growth from above. As such a parameter, we will use the estimate of the bandwidth limit of the communication channel L.

The estimate L is determined empirically for the communication technology used in the investigated information and communication system and the configuration of the hardware component (which in itself is a non-trivial task). In turn, the circular delay is characterized by a stochastic value γ_l, the value of which depends on the size of the sliding window l, which was relevant at the time of transfer of the corresponding data blocks packet. We denote the distribution function of the stochastic value γ_l as $R_l(t)$ and $D_l = \exists\gamma_l$ (if the latter exists).

We focus our research on the description of the information transfer process between the sensor network end IoT device and the hub. The organization of such a process in modern conditions assumes that the monitoring sender (most often – the end IoT device) always has information for transfer. As already mentioned, in real information and communication systems, the bandwidth of the communication channel is limited (estimate L). This circumstance prompts us to introduce a parameter $l_{\max}$ related to the estimate L, the value of which is the upper limit of the growth of the sliding window size.

Based on the above-formulated features of the information transfer process between the sensor network end IoT device and the hub, the analytical description of this process will be carried out based on the stochastic processes theory and the queuing theory. In this context, our goal is to synthesize a queuing system Σ with controlled input flow, deterministic service, feedback, and unlimited queuing. In the terminology of queuing theory:

- a set of data blocks for transfer is a set of requests;
- a communication channel is a service device;
- a service duration distribution is deterministic with a parameter $t_0 = 1/L$.

The functioning of feedback, which regulates the dynamics of the size of the sliding window, is taken into account by entering the set $W = \{W^+, W^-\}$, where W^+ is a service signal about successful data transfer (positive service signal) and W^- characterizes the reversed situation (negative service signal).

The moment of arrival of a signal of type W is a stochastic value: the probability of the appearance of the signal W^+ is characterized by the parameter $1 - p$ and the probability of the appearance of the signal W^- is characterized by the parameter p.

The receipt of requests is regulated by the transport protocol according to the type of service signal W. When a feedback service signal W is received, $k \geq 0$ requests (data blocks) are received from the set of requests to the system Σ. The value k depends on the current size of the sliding window l and the number of facts of receiving negative service signals W^-. Requests available in the system Σ are served sequentially (without regard to priority, in order of arrival). To describe the process of information transfer between the sensor network end IoT device and the hub, it is necessary to analytically characterize the output flow of the system Σ.

We formalize analytically the distribution of the size of the sliding window. Suppose that at the time $t > 0$ the size of the sliding window is determined by the function $l(t)$. Also, let us generalize by the set $\mathrm{T} = \{\tau_i\}$, $i = 1, 2, \ldots$, the sequence of moments when the value of the parameter l changed in response to the service signals W. The development of this concept will be the Markov chain $l_i = l(\tau_i)$, $l_i \in X\left\{\overline{2, l_{\max}}\right\}$, $i = 1, 2, \ldots$; moreover, the minimum size of the sliding window is $l_{\min} = 2$, which corresponds to the specifics of the investigated process. We define the step process $\{l(t)\}_{t>0}$ as semi-Markov. We characterize the event of a transition of chain $\{l_i\}$ from state u to state v in k steps with probability p_{uv}^k, $u, v \in X$. Based on the above, we write: $p_{uv}^k \underset{k\to\infty}{\to} \pi_v$, $\sum\limits_{l=2}^{l_{\max}} \pi_l = 1$.

Let us generalize the set of probabilities $P\{l(t) = l\}$ $P_l(t) = P\{l(t) = l\}$ and conditional mathematical expectations $\alpha_l = A(\tau_{i+1} - \tau_i | l(\tau_i) = l)$.

By definition: $\varepsilon = \tau_{i+1} - \tau_i > 0$ in which case, either $\tau_{i+1} - \tau_i \geq l/L$ or $\tau_{i+1} - \tau_i \geq \gamma_l$. Accordingly, $\exists\varepsilon : F_l(\varepsilon) \leq 1 - \varepsilon$, where $F_l(\varepsilon)$ are the distribution function of the difference $\tau_{i+1} - \tau_i$ and the conditional mathematical expectation α_l exists if D_l exists. Therefore, if a finite mathematical expectation can be determined for a stochastic quantity γ_l then for a stochastic dependence $P_l(t)$ it is possible to write

$$P_l(t) \underset{t\to\infty}{\rightarrow} \alpha_l \pi_l / \sum_{l=2}^{l_{\max}} \alpha_l \pi_l. \tag{2}$$

The logic of reasoning embodied in expression (2) echoes that which is the basis of the ergodic theorem for semi-Markov processes [11].

We define the stationary distribution π_l for the Markov chain $\{l_i\}$ in the form of Chapman's equations:

$$\begin{aligned} \pi_i = f_{i-1}\pi_{i-1} + (1 - f_{2i})\pi_{2i} + \\ +(1 - f_{2i+1})\pi_{2i+1} \forall 2i \leq l_{\max}, \end{aligned} \tag{3}$$

$$\pi_i = f_{i-1}\pi_{i-1} \forall l_{\max} < 2i < 2l_{\max}, \tag{4}$$

where $f_i = (1 - p)^i$ and

$$\pi_{l_{\max}} = f_{l_{\max}-1}\pi_{l_{\max}-1} + f_{l_{\max}}\pi_{l_{\max}}. \tag{5}$$

In the states π_i defined by Equation (3) the system Σ can enter both under the condition of linear growth (under the condition of receiving positive service signals) and under the condition of gradual decline (under the condition of receiving negative service signals).

In the states π_i defined by Equation (4) the system Σ can enter only under the condition of linear growth (provided positive service signals are received).

Equation (5) describes the situation when the system Σ has reached the upper limit of the size of the sliding window $l_{\max}$ and is in this state before the arrival of a negative service signal.

Obtaining an explicit analytical solution to the system of Equations (3) and (4) taking into account Equation (5) is difficult, and its practical implementation will be accompanied by significant computational costs even with the empirical selection of normalizing constants. So, let us resort to the recurrent representation of π_i taking

$$F_i = \prod_{k=1}^{i} f_k,\ j = \lfloor l_{\max}/2 \rfloor. \tag{6}$$

We will obtain:

$$\pi_i = \pi_j C_i, \tag{7}$$

where the values π_j are determined by the normalization condition and

$$C_i = F_{i-1} \forall j < i < l_{\max}, \tag{8}$$

$$C_{i-1} = \frac{1}{f_{i-1}}(C_i - (C_{2i}(1 - f_{2i}) + C_{2i+1}(1 - f_{2i+1}))), \tag{9}$$

$$C_{l_{\max}} = F_{l_{\max}-1}/(1 - f_{l_{\max}}). \tag{10}$$

If we substitute Expressions (8) and (9) into Equation (7) and take into account notation (6) and expression (10), we will obtain the original system of Equations (3), (4) and Expression (5). Therefore, Expression (7) completely determines the distribution of π_l and the recurrent procedure characterized by Expressions (8)–(10) is characterized by linear complexity $O(l_{\max})$.

Let us combine the entities $l(t)$ (the current size of the sliding window) and $n(t)$ (the number of data blocks sent at the size of the sliding window $l(t)$) into a dual function $\eta(t) = \{l(t), n(t)\}$. We accept $n(t) = 1$ each time when, as a result of receiving service signals W, the size of the sliding window $l(t)$ changes. We denote the moment of sending the i-th data block (the moment of completion of service of the i-th request by the system Σ) as τ_i'. Accordingly, if $\tau_{i_1}' > \tau_{i_2}'$ then $i_1 > i_2$. The sequence $\eta_i = \eta(\tau_i')$ is by definition a Markov chain.

Accepting the notation $\pi_\eta = \pi(l,n)$, we formulate the stationary distribution π_η of the Markov chain $\{\eta_i\}$ based on Chapman's equations:

$$\pi(l,n) = (1-p)\pi(l,n-1)\forall l, 1 < n < l, \tag{11}$$

$$\begin{gathered} \pi(l,1) = (1-p)\pi(l-1,l-1) + p\sum_{j=1}^{2l} \pi(2l,j) + \\ + p\sum_{j=1}^{2l+1} \pi(2l+1,j) = 0 \forall 2l \le l_{\max}, \end{gathered} \tag{12}$$

$$\pi(l,1) = (1-p)\pi(l-1,l-1)\forall l_{\max} < 2l < 2l_{\max}, \tag{13}$$

$$\begin{gathered} \pi(l_{\max},1) = (1-p)\pi(l_{\max}-1,l_{\max}-1) + \\ + (1-p)\pi(l_{\max},l_{\max}). \end{gathered} \tag{14}$$

Based on Equation (11), we write

$$\pi(l,n) = (1-p)^{n-1}\pi(l,1) = (1-p)^{n-1}\pi_l. \tag{15}$$

The distribution π_η is determined by Expressions (7) and (15).

Let us focus on defining conditional mathematical expectations α_l. If the system Σ has time to process the requests in the queue before the service signal W arrives, then $\tau_{i+1} - \tau_i = \gamma_l$. Otherwise, $\tau_{i+1} - \tau_i = l/L$. Accordingly:

$$\alpha_l = \int_{lt_0}^{\infty} t dR_l(t) + lt_0 R_l(lt_0),\ t_0 = 1/L. \tag{16}$$

The change in the size of the sliding window carried out by the mechanisms of the transport protocol is one of the main sources of the stochastic nature of the output stream of the system Σ. A significant characteristic parameter of this flow is bandwidth B. The stochastic characteristic B is defined as the ratio of the number of outgoing requests of the system Σ for the time interval $[\tau_i, \tau_{i+1})$ to the duration of this interval.

For the effective application of the transport protocol in wireless communication networks of 5G technology, the analytical formalization of the distribution of the stochastic characteristic B is relevant. Using 5G technology, the authors focus on such an area of its application as eMBB. The Industry 5.0 paradigm declares the active use of virtualization with the effect of presence, which is impossible without the stable transfer of high-resolution video streams. Moreover, it is wireless communication channels that are optimal in terms of expectations of industrialists, officials, environmentalists and consumers.

If the continuous right function $N(t)$ characterizes the number of outgoing requests of the system Σ at the time t then the character B can be described by the expression

$$B = \frac{N(\tau_{i+1}) - N(\tau_i)}{\tau_{i+1} - \tau_i}. \tag{17}$$

Note, that the logic of determining the size of the sliding window implemented in the transport protocol assumes: if $l < \gamma_l$ then $\tau_{i+1} - \tau_i = \gamma$ and $B = l/\gamma_l$. If the condition $l < \gamma_l$ is not fulfilled, then, for an a priori positive queue length of the system Σ, equality $B = L$ holds. We have already obtained the analytical characterization of the independent stochastic value $l(t)$ (the size of the sliding window, see above). Using the full probability formula [19] $\forall 0 \leq x < L$, we define the distribution function $F_B(x)$ as

$$F_B(x) = P(B < x) = \sum_{i=2}^{l_{\max}} P(l = i) P(D > i/x), \quad (18)$$

We introduce the limit P_l defined by expression (2) into expression (18):

$$F_B(x) = \sum_{l=2}^{l_{\max}} P_l (1 - R_l(l/x)) = 1 - \sum_{l=2}^{l_{\max}} P_l R_l(l/x). \quad (19)$$

Expression (19) is valid for $\forall (x < L) \cup (B(L) = 1)$.

Therefore, we have analytically formalized both the distribution of the size of the sliding window and the distribution of the bandwidth of the communication channel for the process of information transfer between the sensor network end IoT device and the hub based on the stochastic processes theory and the queuing theory. The study takes into account that:

- service signals regarding the result of the transfer of a packet of data blocks arrive at random moments, which led to the interpretation of RTT as a stochastic parameter with a known distribution, depending on the size of the sliding window;
- when forming the parametric space of the model of the researched process, it is taken into account that in real information and communication systems both the size of the sliding window and the bandwidth of the communication channel is finite;
- both the sliding window size distribution and the bandwidth distribution are defined in terms of the ratio of RTT and bandwidth of the communication channel. This eliminates the potential influence of the speed characteristics of the data transfer channel on the adequacy of the description of the researched process by the created mathematical apparatus.

4. Results and Discussion

Let us start setting up the experiment by specifying the entities mentioned in the research object, namely, "sensor network ends IoT device" and "hub". By the hub, we understand the data processing center [20] in the classical sense of this term. Now let us define the end IoT device as an integral element of the sensor network. A modern trend in the organization of sensor networks is the use of edge computing technology [21–23] for processing the data collected by the sensors to relieve the burden of the connection channel with the hub. The spread of this technology is explained by an objective fact: the scale of sensor networks is constantly growing. The introduction of edge computing makes it possible to smooth out the problem of controllability of the sensor network in the conditions of its expansion. The implementation of edge computing technology on a cloud computing platform has both advantages (scalability, survivability) and disadvantages (complications of ensuring confidentiality and regular subscription costs).

The authors propose to carry out preprocessing of sensor data on a developed set of low-level data processing centers. Technologically, each such low-level data preprocessing center or end IoT device is a Raspberry Pi computer (the authors focus on the 4B and Compute Module 4 models, which combine a powerful quad-core ARM v8 Cortex-A72 processor with a full range of current communication interfaces, including Gigabit Ethernet, Wi-Fi 802.11ac, Bluetooth 5.0). Another important fact in favor of the Raspberry Pi is that this computer can function under the control of both an open license Linux operating system (individual features of the installation process are solved with the help of the NOOBS tool presented by Raspberry Pi and a licensed Windows 10 IOT operating system. With this interpretation of the sensor network end IoT device, it is possible to implement the process of information transfer both under the control of the TCP protocol (Next Generation TCP/IP Stack) and under the control of the TCP BIC, TCP CUBIC, Highspeed TCP, H-TCP,

TCP Hybla, TCP Illinois, TCP Low Priority, TCP Vegas, TCP NewReno, TCP Veno, TCP Westwood+, YeAH-TCP.

The application of the mathematical apparatus presented in the previous section for estimating the size of the sliding window and the bandwidth of the communication channel is implemented through the possibilities defined by the standards RFC1323, RFC2018, and RFC3168, namely:

- TCP Window Scale Option: the ability to vary the size of the sliding window up to the limit value of $2^{30} = 1$ GB,
- TCP selective acknowledgment (SACK) options: the possibility of feedback (receiving positive and negative sensory signals from the addressee),
- Explicit Congestion Notification (ECN): the ability to detect congestion of the communication channel without losing data packets.
- TCP timestamps: the possibility of more accurate measurement of RTT due to Prevention Against Wrapped Sequence ACK numbers (PAWS).

The external resource https://www.speedguide.net/analyzer.php (accessed on 8 November 2023) was used to check the current settings of the transport protocol.

Other settings of the transport protocol regarding the used operating system were carried out according to the advice of colleagues from the Pittsburgh Supercomputing Center: https://www.psc.edu/research/networking/tcp-tune (accessed on 8 November 2023).

The New TTCP utility (nuttcp: https://www.nuttcp.net/Welcome%20Page.html (accessed on 8 November 2023)) was used to directly measure the performance of the TCP/IP stack. The advantages of this utility are:

- a simple and effective method of measuring bandwidth via TCP or UDP,
- cross-platform,
- the ability to check the effectiveness of the local TCP/IP stack (loopback),
- the correct termination of TCP connections, and the ability to work with clients under NAT.

Let us apply the mathematical apparatus obtained in the previous section, generalized by Expression (19), to the task of analyzing real network information transactions. By default, we assume that the function $R_l(t) = R(t)$ does not depend on the size of the sliding window and is a normal distribution with parameters $\mu = 60$ ms and $\sigma = 30$ ms.

Figure 2 visualizes the stationary distribution of the size of the sliding window π_l for different probabilities p of the appearance of the service signal W^- and at $l_{\max} = 120$ ms (typical value for the Internet): $\pi_l = f(l)$, $p_{W^-}, l_{\max} = \text{const}$.

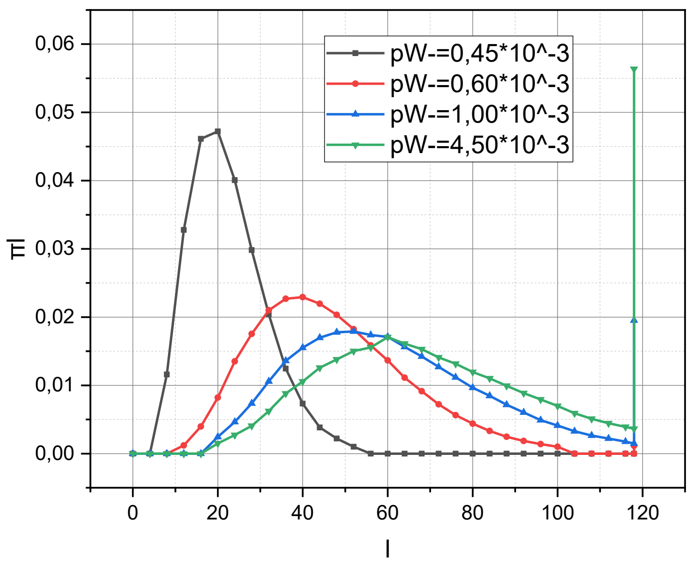


Figure 2. Visualization of the dependence of $\pi_l = f(l)$ at $p_{W^-}, l_{\max} = \text{const}$.

The transport protocol changes the size of the sliding window l according to the ratio between the maximum size of the sliding window $l_{\max}$ and the probability of losing a packet of data blocks p_{W^-}. It is possible to track three main modes in the dynamics of changing the value of the size of the sliding window. In the first mode, the size of the sliding window quickly reaches its maximum value and keeps it almost all the time during the information transaction. In the second mode, the arrival of a negative service signal W^- leads to a reduction in the size of the sliding window by half with a quick return to the maximum value $l_{\max}$. In this mode, the corresponding distribution π_l has two extremes at the points $l_{\max}$ and $l_{\max}/2$. As the probability of the appearance of a negative service signal p_{W^-} increases, the duration of the sliding window size in a stable state (third mode) decreases. The distribution π_l loses its extremum at this point $l_{\max}/2$. If the probability of the appearance of a negative service signal W^- exceeds 0.1, then the value $l_{\max} > 10$ does not have a noticeable effect on the appearance of the distribution π_l. Breaks of the smooth nature of individual curves from Figure 2 are explained by the chosen mathematical apparatus for describing the studied process which is a priori stepwise.

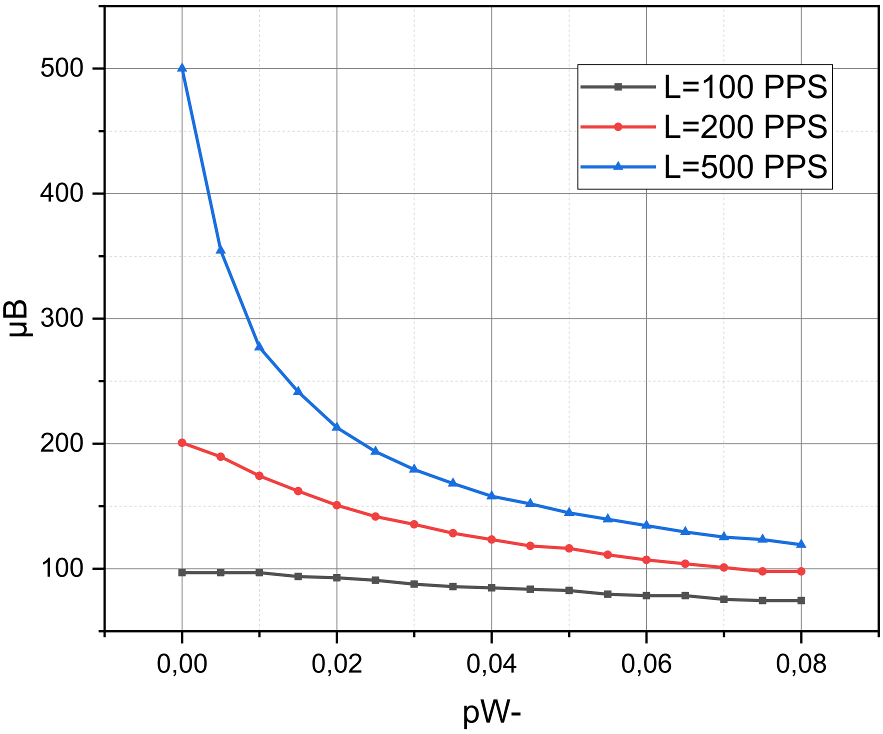
Figure 3 shows the mathematical expectation μ_B and the root mean square deviation σ_B of the bandwidth B as a function of the probability p_{W^-}. The curves are obtained for different values of the estimate L at a constant value of $l_{\max} = 70$.

From Figure 3a it can be seen that the curves of the mathematical expectation μ_B are monotonic decreasing functions under conditions of the increasing value of the argument p_{W^-}. Such a theoretically determined trend coincides with the real behavior of the network construct, which was investigated by the nuttcp utility. It can be seen that the corresponding value of the estimate L determines the initial value of the curve $\mu_B = f(p_{W^-})$ and determines the rate of its decreasing with increasing probability p_{W^-}. Thus, the accuracy of the estimation of the bandwidth of the communication channel is an important factor that affects the performance of information transfer in conditions of low interference (low probability of the appearance of a negative service signal W^-). This is an important conclusion because it is a factor that should determine the shift in the focus of attention of researchers from the search for more accurate methods of estimating the bandwidth of a communication channel (low probability p_{W^-}) to searching for more secure methods of encoding data packets or causes of interference (high probability p_{W^-}).

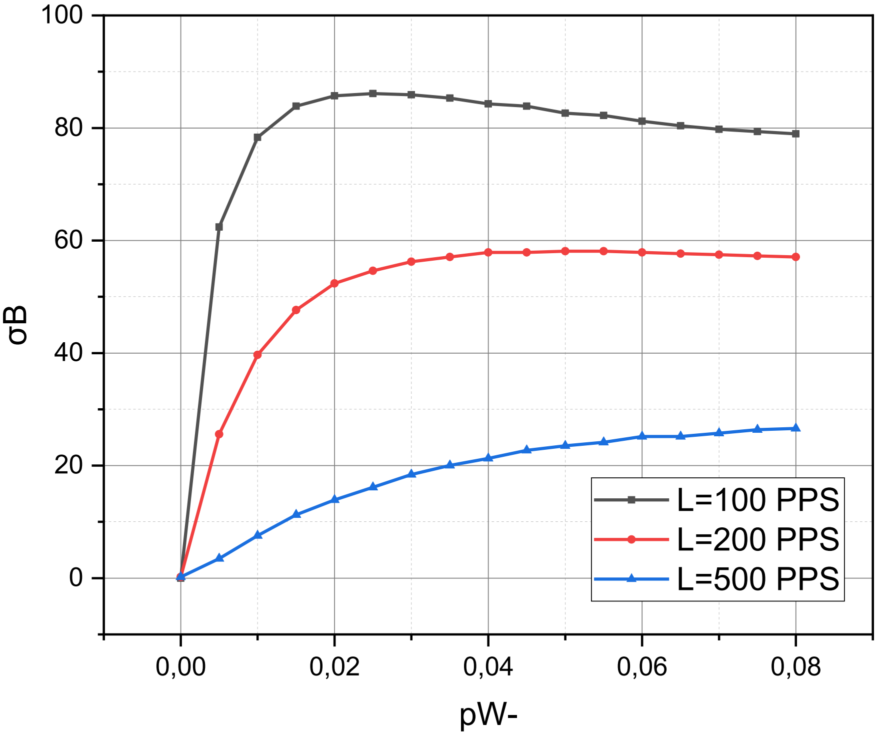
Figure 3b shows that the function $\sigma_B = f(p_{W^-})$ can be both monotonic and have an extremum on the interval $p_{W^-} \in (0, 1)$. This fact allows us to predict the possibility of setting the problem of finding the optimal value of the bandwidth of the communication channel for the parametric space L, $l_{\max}$, p_{W^-}.

Figure 4 visualizes the mathematical expectation μ_B and the root mean square deviation σ_B of the bandwidth B as a function of the circular delay parameters D. The arguments are the probability values p_{W^-}.

Note, that the nature of the functional dependencies presented in Figures 3 and 4, coincides, which indicates the relationship between the parameters L D. Figure 4 shows that the curves of the mathematical expectation μ_B are monotonic decreasing functions under conditions of the increasing value of the argument p_{W^-}. Such a theoretically determined trend coincides with the real behavior of the network construct, which was investigated by the nuttcp utility. It is interesting that the curve σ_B can be both monotonic and have an extremum on the interval $p_{W^-} \in (0, 1)$. This circumstance makes it possible to supplement the parametric space of the potential problem of finding the optimal value of the bandwidth of the communication channel with the parameter D.



(**a**)



(**b**)

Figure 3. (**a**) Visualization of the dependence of $\mu_B = f(p_{W-})$ at $L =$ const. (**b**) Visualization of the dependence of $\sigma_B = f(p_{W-})$ at $L =$ const.

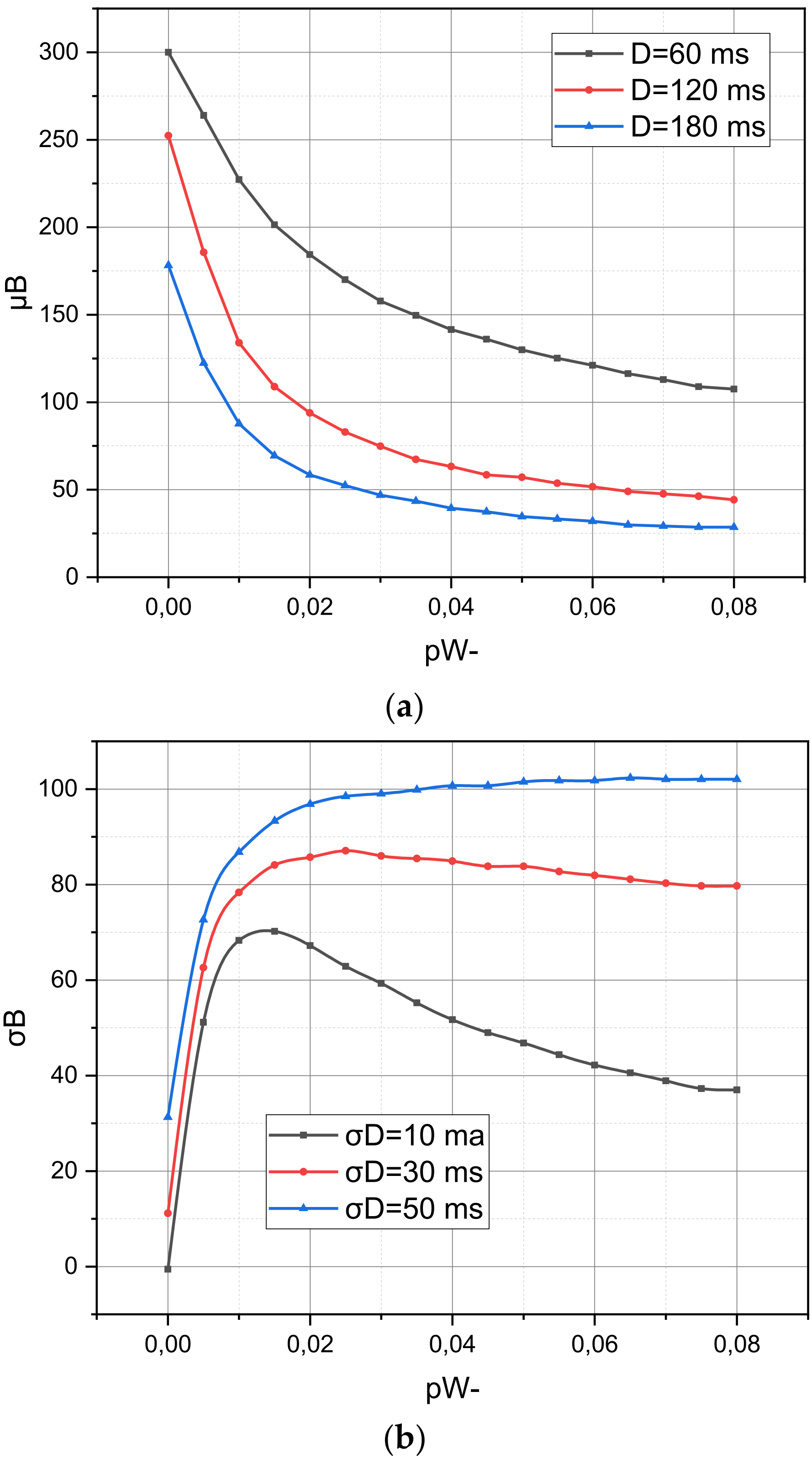


Figure 4. (**a**) Visualization of the dependence of $\mu_B = f(p_{W^-})$ at $D = \text{const}$. (**b**) Visualization of the dependence of $\sigma_B = f(p_{W^-})$ at $\sigma_D = \text{const}$.

We will conclude the experimental section by comparing the TCP Westwood+ protocol, the Window Scale parameter of which was determined based on the author's mathematical apparatus, with the BIC TCP, TCP Vegas, TCP NewReno, TCP Veno protocols without tuning (see Figure 5). The communication channel was supported by 5G technology.

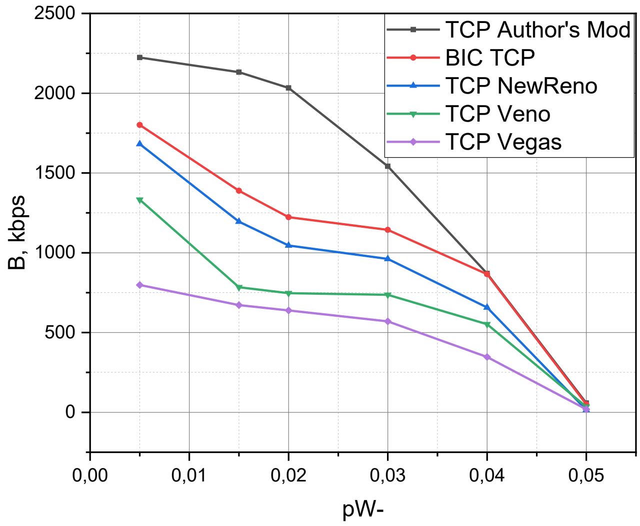


Figure 5. Empirical dependences $B = f(p_{W^-})$ for a stable wireless network configuration and different transport protocols.

The results shown in Figures 3 and 4 confirm the adequacy of the mathematical apparatus proposed by the authors, and the results shown in Figure 5 testify to its applied potential. However, this potential can be fully revealed only by researching the relevant optimization problems.

5. Conclusions

The object of this research is the transport layer for managing the process of data transfer between the sensor network end IoT device and the hub using the communication capabilities of the 5G platform. In this case, the authors focus their attention on the TCP protocol. We interpreted the research object as a semi-Markov (focused on the dynamics of the size of the sliding window of the protocol) process with two nested Markov chains (the first characterizes the current size of the sliding window, and the second—the number of data blocks sent at the current value of this characteristic).

As a result, a stationary distribution of the size of the sliding window (a parameter that determines the intensity of the information flow from the addressee) was obtained both for the resulting semi-Markov process and for nested Markov chains, etc. A recursive approach to the calculation of the mentioned stationary distribution, which is characterized by linear computational complexity, is formalized. Based on the obtained stationary distribution of the size of the sliding window, a distribution function is formulated that characterizes the bandwidth of the communication channel between the entities specified in the researched process.

Future research. As we mentioned earlier, the object of our research is the transport layer for managing the process of data transfer between the sensor network end IoT device and the hub using the communication capabilities of the 5G platform. In this article, we presented the basic mathematical apparatus for studying this process. We see its further application in determining the optimal TCP parameters for exact QoS policies that will be used to manage information interaction in a 5G cluster that supports the operation of a sensor network using URLLC, mMTC, and eMBB technologies.

The results presented in the article showed the promise of using the TCP protocol in a sensor network, the end IoT devices of which generate massive traffic. At the same time, we note that the TCP protocol has some problems with data security, namely [24–26]:

- TCP is incapable of safeguarding a segment against message modification attacks due to its lack of protection for the checksum field. This field is intended to detect alterations in a segment, but it remains vulnerable to message modification attacks, allowing for the manipulation of TCP segments without detection. Additionally, there are no mechanisms for peer entities to detect message modification attacks.
- TCP does not provide data encryption capabilities, making it unable to maintain the security of segment data against message eavesdropping attacks. TCP transports unencrypted data from the application layer, leaving any valuable information exposed to potential interception.
- TCP is unable to defend connections against unauthorized access attacks because it verifies a peer entity solely based on the source IP address and port number, which can be easily modified by attackers.

We will try to remove these limitations in our future studies.

Author Contributions: Conceptualization, V.K.; methodology, V.K.; software, V.K.; validation, K.G. and K.P.; formal analysis, V.K.; investigation, V.K.; resources, K.G. and K.P.; data curation, K.G. and K.P.; writing—original draft preparation, V.K.; writing—review and editing, V.K.; visualization, V.K.; supervision, V.K.; project administration, V.K.; funding acquisition, V.K. All authors have read and agreed to the published version of the manuscript.

Funding: Project "Methodology for Increasing the Dependability of Information Systems for Critical Use with a Heterogeneous Wireless Interface", reg. no. 2022/45/P/ST7/03450, the POLONEZ BIS 2 program, implemented by the National Science Center in Krakow.

Data Availability Statement: Most data is contained within the article. All the data are available on request due to restrictions, e.g., privacy or ethics.

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

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kovtun, V.; Grochla, K. Investigation of the Competitive Nature of eMBB and mMTC 5G Services in Conditions of Limited Communication Resource. *Sci. Rep.* **2022**, *12*, 16050. [CrossRef] [PubMed]
2. Kovtun, V.; Izonin, I.; Gregus, M. Formalization of the Metric of Parameters for Quality Evaluation of the Subject-System Interaction Session in the 5G-IoT Ecosystem. *Alex. Eng. J.* **2022**, *61*, 7941–7952. [CrossRef]
3. Bilyk, O.; Obelovska, K. Power Consumption Analysis at MAC-Sublayer of Wireless Sensor Networks. In *Advances in Artificial Systems for Logistics Engineering*; Springer: Cham, Switzerland, 2022; pp. 27–36. [CrossRef]
4. Lin, J.; Cui, L.; Zhang, Y.; Tso, F.P.; Guan, Q. Extensive Evaluation on the Performance and Behaviour of TCP Congestion Control Protocols under Varied Network Scenarios. *Comput. Netw.* **2019**, *163*, 106872. [CrossRef]
5. Marin, A.; Rossi, S.; Zen, C. Size-Based Scheduling for TCP Flows: Implementation and Performance Evaluation. *Comput. Netw.* **2020**, *183*, 107574. [CrossRef]
6. Kovtun, V.; Altameem, T.; Al-Maitah, M.; Kempa, W. The Markov Concept of the Energy Efficiency Assessment of the Edge Computing Infrastructure Peripheral Server Functioning over Time. *Electronics* **2023**, *12*, 4320. [CrossRef]
7. Ding, L.; Tian, Y.; Liu, T.; Wei, Z.; Zhang, X. Understanding Commercial 5G and Its Implications to (Multipath) TCP. *Comput. Netw.* **2021**, *198*, 108401. [CrossRef]
8. Bruhn, P.; Kühlewind, M.; Muehleisen, M. Performance and Improvements of TCP CUBIC in Low-Delay Cellular Networks. *Comput. Netw.* **2023**, *224*, 109609. [CrossRef]
9. Li, F.; Guo, Z.; Liang, B.; Yi, X.; Wang, X.; Li, W.; Wang, Y. A Measurement Study on Device-to-Device Communication Technologies for IIoT. *Comput. Netw.* **2021**, *192*, 108072. [CrossRef]
10. Kovtun, V.; Altameem, T.; Al-Maitah, M.; Kempa, W. Information Technology for Maximizing Energy Consumption for Useful Information Traffic in a Dense Wi-Fi 6/6E Ecosystem. *Electronics* **2023**, *12*, 3847. [CrossRef]
11. Hurni, P.; Bürgi, U.; Anwander, M.; Braun, T. TCP Performance Optimizations for Wireless Sensor Networks. In Proceedings of the Wireless Sensor Networks: 9th European Conference, EWSN 2012, Trento, Italy, 15–17 February 2012; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7158, pp. 17–32. [CrossRef]
12. Kim, H.-S.; Im, H.; Lee, M.-S.; Paek, J.; Bahk, S. A Measurement Study of TCP over RPL in Low-Power and Lossy Networks. *J. Commun. Netw.* **2015**, *17*, 647–655. [CrossRef]

13. Park, M.; Paek, J. TAiM: TCP Assistant-in-the-Middle for Multihop Low-Power and Lossy Networks in IoT. *J. Commun. Netw.* **2019**, *21*, 192–199. [CrossRef]
14. Gomez, C.; Arcia-Moret, A.; Crowcroft, J. TCP in the Internet of Things: From Ostracism to Prominence. *IEEE Internet Comput.* **2018**, *22*, 29–41. [CrossRef]
15. Kumar, S.; Michael, P.A.; Kim, H.-S.; Culler, D.E. Tcplp: System design and analysis of full-scale TCP in low-power networks. *arXiv* **2018**, arXiv:1811.02721. Available online: https://www.researchgate.net/profile/Hyung-Sin-Kim/publication/328800924_TCPlp_System_Design_and_Analysis_of_Full-Scale_TCP_in_Low-Power_Networks/links/5bef0defa6fdcc3a8ddbdd21/TCPlp-System-Design-and-Analysis-of-Full-Scale-TCP-in-Low-Power-Networks.pdf?origin=publicationDetail&_sg%5B0%5D=ao08rSTIp-huC9mwHDffX6nPqYWS-F-bllTl6AwYmksypELboerDfT-s8mOKnN6ukMDVtaoguaDN3cHFBjsmkw.NoAQuyppjva5rfmHUWLDodjPT7xtjCGFkUuOig7fk5BY3e-WSlFoWX-YyKmIHwc6-hqfBHzNcfGzhz0fR2kp-w&_sg%5B1%5D=9MFgJwBY3EWoM_n8znysdSBewA54Si3mSzqlmTi8SXdZ45TPwdNyFsnTsE88RUd66bFoFJJFCKp9TX9BMr6jWC0rQn-RWY8Rk3TIkPzOMqHi.NoAQuyppjva5rfmHUWLDodjPT7xtjCGFkUuOig7fk5BY3e-WSlFoWX-YyKmIHwc6-hqfBHzNcfGzhz0fR2kp-w&_iepl=&_rtd=eyJjb250ZW50SW50ZW50IjoibWFpbkl0ZW0ifQ%3D%3D&_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCIsInBvc2l0aW9uIjoicGFnZUhlYWRlciJ9fQ (accessed on 8 November 2023).
16. Obelovska, K.; Snaichuk, Y.; Selecky, J.; Liskevych, R.; Valkova, T. An Approach Toward Packet Routing in the OSPF-Based Network with a Distrustful Router. *Wseas Trans. Inf. Sci. Appl.* **2023**, *20*, 432–443.
17. Auzinger, W.; Obelovska, K.; Dronyuk, I.; Pelekh, K.; Stolyarchuk, R. A Continuous Model for States in CSMA/CA-Based Wireless Local Networks Derived from State Transition Diagrams. *Proc. Int. Conf. Data Sci. Appl.* **2021**, 571–579. [CrossRef]
18. Izonin, I.; Tkachenko, R.; Krak, I.; Berezsky, O.; Shevchuk, I.; Shandilya, S.K. A Cascade Ensemble-Learning Model for the Deployment at the Edge: Case on Missing IoT Data Recovery in Environmental Monitoring Systems. *Front. Environ. Sci.* **2023**, *11*, 1295526. [CrossRef]
19. Semenov, A.; Semenova, O.; Kryvinska, N.; Tromsyuk, V.; Tsyrulnyk, S.; Rudyk, A.; Kacprzyk, J. Advanced Correlation Method for Bit Position Detection towards High Accuracy Data Processing in Industrial Computer Systems. *Inf. Sci.* **2023**, *624*, 652–673. [CrossRef]
20. Kor, A.-L.; Yanovsky, M.; Pattinson, C.; Kharchenko, V. SMART-ITEM: IoT-Enabled Smart Living. In Proceedings of the 2016 Future Technologies Conference (FTC), San Francisco, CA, USA, 6–7 December 2016. [CrossRef]
21. Zaitseva, E.; Rabcan, J.; Levashenko, V.; Kvassay, M. Importance Analysis of Decision Making Factors Based on Fuzzy Decision Trees. *Appl. Soft Comput.* **2023**, *134*, 109988. [CrossRef]
22. Durnyak, B.; Tymchenko, B.H.O.; Tymchenko, O.; Anastasiya, D. Research of Image Processing Methods in Publishing Output Systems. In Proceedings of the 2018 XIV-th International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), Lviv, Ukraine, 18–22 April 2018. [CrossRef]
23. Mochurad, L.I. Canny Edge Detection Analysis Based on Parallel Algorithm, Constructed Complexity Scale and CUDA. *Comput. Inform.* **2022**, *41*, 957–980. [CrossRef]
24. Kou, L.; Wu, J.; Zhang, F.; Ji, P.; Ke, W.; Wan, J.; Liu, H.; Li, Y.; Yuan, Q. Image Encryption for Offshore Wind Power Based on 2D-LCLM and Zhou Yi Eight Trigrams. *International J. Bio-Inspired Comput.* **2023**, *22*, 53–64. [CrossRef]
25. Cheng, Y.; Liu, Y.; Zhang, Z.; Li, Y. An Asymmetric Encryption-Based Key Distribution Method for Wireless Sensor Networks. *Sensors* **2023**, *23*, 6460. [CrossRef] [PubMed]
26. Hazaimeh, O.M.A.; Al Jamal, M.F.; Alomari, A.; Bawaneh, M.J.; Tahat, N. Image Encryption Using Anti-Synchronisation and Bogdanov Transformation Map. *Int. J. Comput. Sci. Math.* **2022**, *15*, 43. [CrossRef]