

## Article

# A Multigraph-Defined Distribution Function in a Simulation Model of a Communication Network

Slobodan Miletic <sup>1</sup>, Ivan Pokrajac <sup>1</sup>, Karelia Pena-Pena <sup>2</sup> , Gonzalo R. Arce <sup>2</sup>  and Vladimir Mladenovic <sup>3,\*</sup> <sup>1</sup> Electronic Systems Department, Military Technical Institute, 11000 Belgrade, Serbia<sup>2</sup> Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA<sup>3</sup> Faculty of Technical Sciences Cacak, University of Kragujevac, 34000 Kragujevac, Serbia

\* Correspondence: vladimir.mladenovic@ftn.kg.ac.rs

**Abstract:** We presented a method based on multigraphs to mathematically define a distribution function in time for the generation of data exchange in a special-purpose communication network. This is needed for the modeling and design of communication networks (CNs) consisting of integrated telecommunications and computer networks (ITCN). Simulation models require a precise definition of network traffic communication. An additional problem for describing the network traffic in simulation models is the mathematical model of data distribution, according to which the generation and exchange of certain types and quantities of data are realized. The application of multigraphs enabled the time and quantity of the data distribution to be displayed as operational procedures for a special-purpose communication unit. A multigraph was formed for each data-exchange time and allowed its associated adjacency matrix to be defined. Using the matrix estimation method allowed the mathematical definition of the distribution function values. The application of the described method for the use of multigraphs enabled a more accurate mathematical description of real traffic in communication networks.

**Keywords:** communication network; multigraphs; adjacency matrix; network simulation; network traffic; distribution function



**Citation:** Miletic, S.; Pokrajac, I.; Pena-Pena, K.; Arce, G.R.; Mladenovic, V. A Multigraph-Defined Distribution Function in a Simulation Model of a Communication Network. *Entropy* **2022**, *24*, 1294. <https://doi.org/10.3390/e24091294>

Academic Editor: Raúl Alcaraz

Received: 2 June 2022

Accepted: 5 September 2022

Published: 14 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The design of communication networks as a spatially distributed integrated telecommunication and computer network (ITCN) has been improved by the application of computer simulations. Defining a simulation model of an ITCN is realized by using advanced simulation software with integrated tools. These tools allow an analysis of the network elements' parameters. The application of this methodology of ITCN simulation model design requires a precise definition of network traffic in addition to a definition of the active and passive elements of the architecture and network topology [1]. Network traffic is the process of time events generating a certain type and amount of data at a source, and their distribution between sources and destinations connected in a communication network. There may be multiple data sources in a network that generate the same or different types and amounts of data at the same or different times, and destinations may simultaneously receive data from one or more sources. The problem in designing a simulation model is generating an accurate description of this network traffic. Depending on the purpose of the communication network, the network traffic process can be a deterministic or stochastic event.

In accordance with this definition, network traffic requires the application of appropriate models and distribution functions of communication data over time. An additional problem of describing network traffic in the simulation model is defining the mathematical model according to which the generation of and change in the amount of data are realized. The model requires an appropriate distribution function that, in the simulation model, temporally describes the generation and distribution of the amount of data between

network elements. We used the sampling matrix associated with multigraphs [2] to derive the time distribution function of the communication events of network traffic. A method of applying multigraphs for defining the distribution functions of the generation time of data between network elements of the ITCN is presented in this study. The contributions of this study are as follows:

- A new method of defining network traffic was proposed. The distribution function for creating a simulation model of a communication network was developed, based on the description of communication events and the values of the parameters they determined. The application of this method enabled us to solve the problem of describing the time of data generation and distribution in the communication networks.
- The application of multigraphs for the mathematical derivation of a more precise distribution function of data was proposed and compared with other methods in which the distribution function of data was approximated by the type of network traffic and by the time variation of the data.
- The application of multigraphs and their related matrices enabled multiple descriptions of network traffic in terms of events and communication parameters, which enabled their change in time to be mathematically represented as a function of the schedule. The new approach enabled a more accurate description of the network traffic in the design of a simulation model of the communication network and time-accurate results in the simulation.

The paper is organized as follows. Section 2 provides an overview of the different methods used for defining and statistically describing network traffic. Definitions of all the starting elements needed to describe network communications are presented in Section 3. In Section 4, the basic concept, and details of the proposed method of applying multigraphs for describing the time of network traffic and data distribution are presented. Section 5 presents an application of the mathematical derivation, and a graphical representation of the time distribution functions in the proposed method. Section 6 concludes the study and gives directions for further research in the application of the method.

## 2. Related Work

In earlier works, different methods of defining network traffic were proposed. Network traffic is a complex time-stochastic or -deterministic process of network structure. Earlier methods consisted of complex procedures for describing and defining the network traffic. The methods are complex, especially for describing network traffic with the distribution of multiple data formats. The basic method of defining network traffic was realized by measuring and recording traffic in test networks, as in [3–7], with the theoretical derivation of statistical mathematical descriptions for further analysis. Measurements of the generated and distributed network traffic enabled statistical descriptions and parametric descriptions in [4,5,8]. In recent work [3,6,7,9–13], statistical typification of network traffic with known distribution functions and parameter variation was defined or the traffic was described by using self-similarity related to heavy-tail distributions [14,15]. Further definitions of network traffic were limited to the recognition network traffic type (voice internet, HTTP, VoIP, multimedia, etc.) and descriptions of the intended distribution function. An overview and comparison of the methods used in previous studies to define network traffic are given in Table 1.

**Table 1.** Review of the network traffic defining methods.

Reference	Methods	Measurement Source	Statistical Description	Traffic	Illustrating	Application	Country	Year
[3]	Traffic self-similarity, the approximation function of traffic	The average daily traffic recorded	Pareto distribution	2G, voice, HSDPA,	Function distribution graph	Simulating real network traffic	Russia	2021
[4]	Nonlinear analysis of traffic measurements	A medium-sized LAN with 200 to 250 interconnected computers	Kolmogorov's scheme for describing network traffic, log-normal distribution, Gaussian distribution	NetBEUI, TCP/IP	Function distribution graph	Realistic dynamical models of network traffic	Russia	2004
[5]	Mathematical approximation	Traffic volume recorded by routers, ethernet traffic traces	Poisson's probability distribution	Ethernet, MPEG4, TCP/IP, web, email, multimedia	Function distribution graph	Traffic modeling	USA, Texas	2007
[6]	Self-similarity statistical analysis of network traffic measurements	Computer network in small company	Gaussian or power-law probability distributions	Web, HTTP, internet, email, SSL, IPv6	Function distribution graph	Computer network traffic analysis	Poland	2021
[7]	Statistical analysis	Academic, commercial and residential networks; data centers	Log-normal distribution, Gaussian distribution, Weibull distribution	Internet IPv4	Function distribution graph	Predicting the proportion of time traffic, statistically predicted outcomes for the network	USA, Chicago	2019
[8]	Introductory techniques for input modeling; graphical and statistical methods; mathematics	Sample statistics, the Kolmogorov–Smirnov test statistic, the discrete-event simulation, hypothetical arrival process, stochastic processes	Binomial, degenerate Normal, exponential, Bezier curve, independent binomial, bivariate exponential, Markov chain, Poisson process, nonhomogeneous Poisson process, Markov process	Discrete, continuous modeling arrivals	Histogram, function distribution graph	Input models available to simulation analysts	USA	2001

Table 1. Cont.

Reference	Methods	Measurement Source	Statistical Description	Traffic	Illustrating	Application	Country	Year
[9]	Simulation modeling process	Describing the behaviors and interactions	Classical statistics right-triangular distribution, cumulative distribution function, uniform distribution	Discrete event systems		Simulating and modeling operations, distribution modeling	USA	2007
[16]	Traffic modeling	Core router of a university, backbone links trans-Pacific backbone link	Gaussian distribution	Gaussian traffic model	Q–Q plots, timescales	Network modeling	Netherlands, Denmark	2013
[10]	Traffic analysis	Counting process, inter-arrival time process, discrete-time traffic,	Poisson, Pareto, Weibull, Markov, Markov chain, on–off model, interrupted Poisson	The traffic on the network	Mathematically, graphs	Traffic modeling, capacity planning the design of networks and services		
[17]	Traffic analysis	The University of Jordan’s network	Poisson traffic model, long-tail traffic models	Internet traffic	Daily traffic flow graph	Traffic model QoS	Jordan	2019
[11]	Traffic analysis, mathematics	1998 FIFA World Cup website	Poisson traffic model, Gaussian distribution	Internet traffic	Function distribution graph	Simulation model	Russia	2017
[12]	Traffic analysis	Hubs of cities in Europe and America	Normal probabilistic	Internet traffic, web traffic	Traffic flow graph, probabilistic distribution	Simulation model	Ukraine	2019
[13]	Traffic analysis	Computer network traffic		Multimedia, VoIP	Average daily computer network traffic	Network modeling	Ukraine	2019

These methods are integrated into simulation tools such as OPNET and other advanced software simulation packages. The application of the methods described in previous research may lead to incorrect selection or description of the statistical parameters of the distribution function. The consequence is that one may obtain incorrect simulation results and derive erroneous conclusions and decisions about the design of the network structure. To increase the accuracy of parameters when describing and defining network traffic in [2], we performed an analytical method where we used multigraphs to describe communication interactions as events between network elements. This method was executed and tested on the example of deterministic arranged communication in the network.

The network traffic matrix model has a significant role in network design, network traffic design and analysis of the results, as in the method given in [18,19]. In the method proposed in this study, we introduced a new approach to defining the basis required to obtain a mathematical model from the network traffic matrix. By applying the mathematical models given in [20,21], the time dimension of the multigraph was added and the estimated distribution function that describes the network traffic as a statistical time event can be obtained.

### 3. Data Exchange in the Communication Network

Central to definition of network traffic models is the matrix of network traffic between the source and the destination of communications in the network. To achieve functional relations between the participants in communication, the type of necessary communication is defined for which information flows are determined. The realization and establishment of information flows in an ITCN require the application of appropriate network application services, marked as (S1, S2, Sn). The data exchange and information are the basis for defining the moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  when the participants in communication establish their communication interaction, achieve mutual communication, and, at the same time, exchange certain types and amounts of data.

The moments of time in which communication should take place, the duration of the communication, and the types of information provided for the data exchanged in communication are the particular operational procedures of communication. In Figure 1, the operational procedures determine the moments of change in the amount of data to be distributed between elements of the organizational structure. The type of data and the information (voice, message, symbol, text, table, image, video, etc.) to be exchanged between participants in the communication process are defined by the operational procedures for the command operational function. Starting from the defined communication relationships between the elements of the organization, the matrix of network traffic can be obtained.

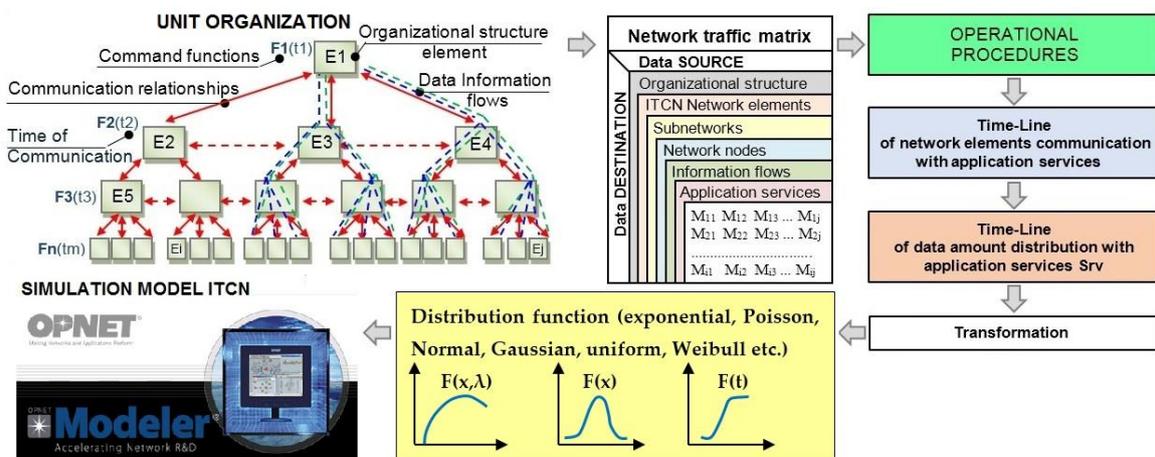


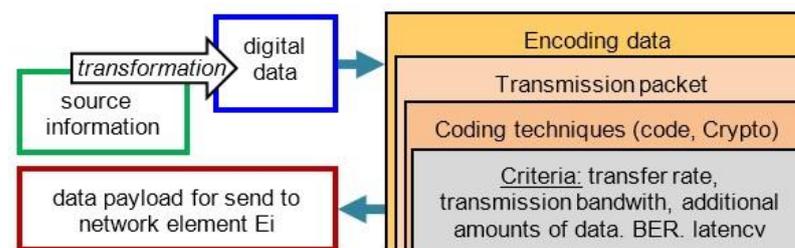
Figure 1. The basic concept of the mapping timeline of the network elements' distribution of data in the OPNET simulation model.

In accordance with the concepts given in [2], communication data exchanges are defined. The time distribution of the data is mapped to the distribution function for the OPNET simulation model, which is shown by the logical flow in Figure 1.

### 3.1. The Data of Network Distribution over Time

Designing a simulation model required us to define the variation in the amount of data generated at the source sent to the destination. The value of the amount of data (Adt) in kbps or Mbps was determined for each type of information (voice, message, symbol, text, table, database, image, video, etc.) exchanged between the elements.

The distribution of the amount of data occurs at the moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ . Additionally, depending on the function in the communication process, the minimum and maximum amount of data generated by the network element,  $E_i$ , for distribution in the network is calculated. The transfer of information to the ITCN requires encrypting of the communication channels. The amount of data for distribution in the network can be increased by the amount of digital code required for protecting information (reconstruction, encryption, error detection). The increase in the amount of data is realized in relation to the header size of the individual layers of the OSI network model. The choice of the access technique, the technology and transmission medium, the communication protocols, and the data packet size (MTU) affect the amount of data to be transferred by the telecommunication links in the ITCN. The total data payload for distribution by the network from the source to destination is determined by the steps in the procedure shown in Figure 2.



**Figure 2.** The procedure for determining the amount of data to send to the network elements in the ITCN.

The distribution function of communication interactions between network elements according to the method given in [2] represents the law of data generation over time. Data generation is realized by the application services.

### 3.2. Distribution Function for Variations in the Amount of Data

The accuracy of the network traffic simulation results in the OPNET simulation model is conditioned by the choice of the distribution function. The distribution function should describe the generation of data and the variation in the amount of data over time. The selection of the distribution function requires one to define its parameters. Moreover, the distribution function is based on the statistical study of communication in the network traffic record, as in [3–7,11–13], or is based on an approximation of the type of network traffic (audio, messages, text, IP, VoIP, video, HTTP, web, ATM, etc.) with existing known distribution functions (exponential, Poisson, Normal (Gaussian), uniform, Weibull etc.), as in [8,10,16,17]. Data traffic modeling is based on self-similarity with the Pareto distributions and the  $\alpha$ -stable distributions, as in [14,15]. Two basic parameters describe the event of data generation at the source in the simulation model. The data generation time is the first parameter. The time of data generation should be adjusted by the time of establishing communication between the network elements. The variation in the amount of data over the duration of the communication is the second parameter. If one chooses an inappropriate distribution function, or by incorrectly defining the value of the variation in the amount of data, or by incorrectly defining the time, the network traffic will be described incorrectly. Simulations of incorrectly described network traffic will not match the predicted network

traffic in an ITCN. In that case, the simulation results are not accurate for analyzing and optimizing the communication network. The error in describing the amount of data generated over time is reduced when the amount of data generated is defined based on the ITCN network traffic matrix and the corresponding distribution function.

#### 4. Description of the ITCN Network Distribution Using Multigraphs

The network data distribution in an ITCN is realized based on the operating procedures and according to the methodology specified in [1]. The information flows of the distribution of the predicted types of information between the network elements are described as well. The multigraphs are defined based on this description. The application of multigraphs allows the time relationships of data distribution to be displayed based on the operational communication procedures. For each moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  when data are exchanged, a multigraph is formed. The formed multigraph is joined with the similarity matrix. The corresponding value of the distribution function for each moment of time is calculated by mathematical estimation of the similarity matrix associated with the multigraph. The use of all the calculated values for all moments of time in the communication interval  $\Delta T = [t_0, t_m]$  enables the definition of the appropriate distribution function.

##### 4.1. Data Distribution Time Scheme between ITCN Network Elements

The generation and distribution of data between network elements in the ITCN are realized through the network application services  $Srv, rv = (1, 2, \dots)$ . Each application service in the ITCN enables the establishment of communication and network distribution of the appropriate type of data. The  $Srv$  application service on the network element  $E_i$  is activated by establishing a communication interaction between the network elements  $E_i$  and  $E_j$  ( $i \neq j$ ) at the moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ . The moments of time are set at the beginning of the time interval in which the application service is active between the network elements. The generation of communication information is enabled and transformed into the appropriate amount of digital data for distribution to the network element  $E_j$  is performed. The time scheme of communication interactions (Figure 3), as in [1,2], shows the flow of these activities from the operational procedures.

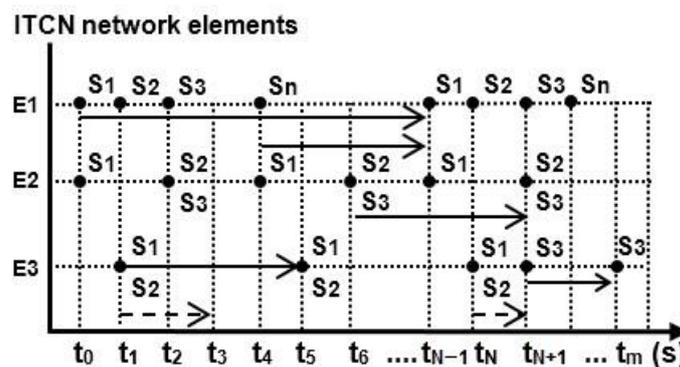


Figure 3. Timeline of activation and repetition of the network elements' communication interactions and the network application services ( $Srv$ ).

The individual timeline of the individual service  $Srv$  now of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  of activation are separated from the given time scheme. The amount of data  $Adt = (Adt_0, Adt_1, \dots, Adt_{N-1}, Adt_N, Adt_m)$  generated at the moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  in the application service  $Srv$  is also defined. Examples of the separate individual timing schemes for Services  $S_1$  and  $S_2$  are displayed in Figure 4.

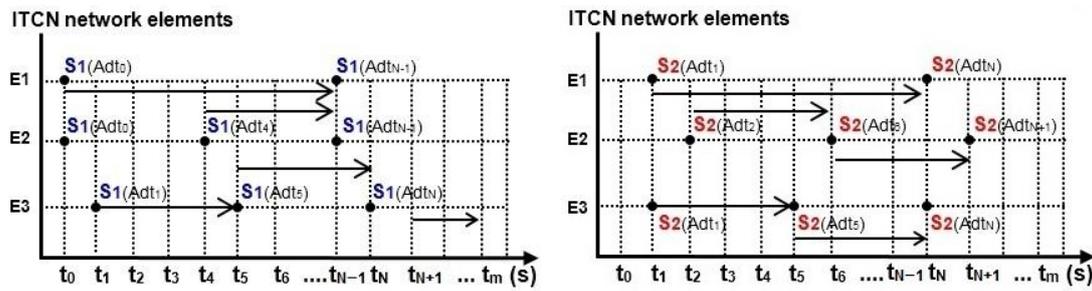


Figure 4. Timeline of activation and repetition of network application services S1 and S2 with amounts of generated data Adt.

Other ways of representing the communication activation of application services S1 to S4 between network elements E1 to E8 are shown in Figure 5. This representation is used to define the amount of data generated Adt at the moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ . For example, the service S1 in the network element E1 with a data quantity of  $Adt_0 = 15$  kbps for distribution to the network element E2 at time  $t_0$  is denoted as E1E2S1\_Ad $t_0$ .

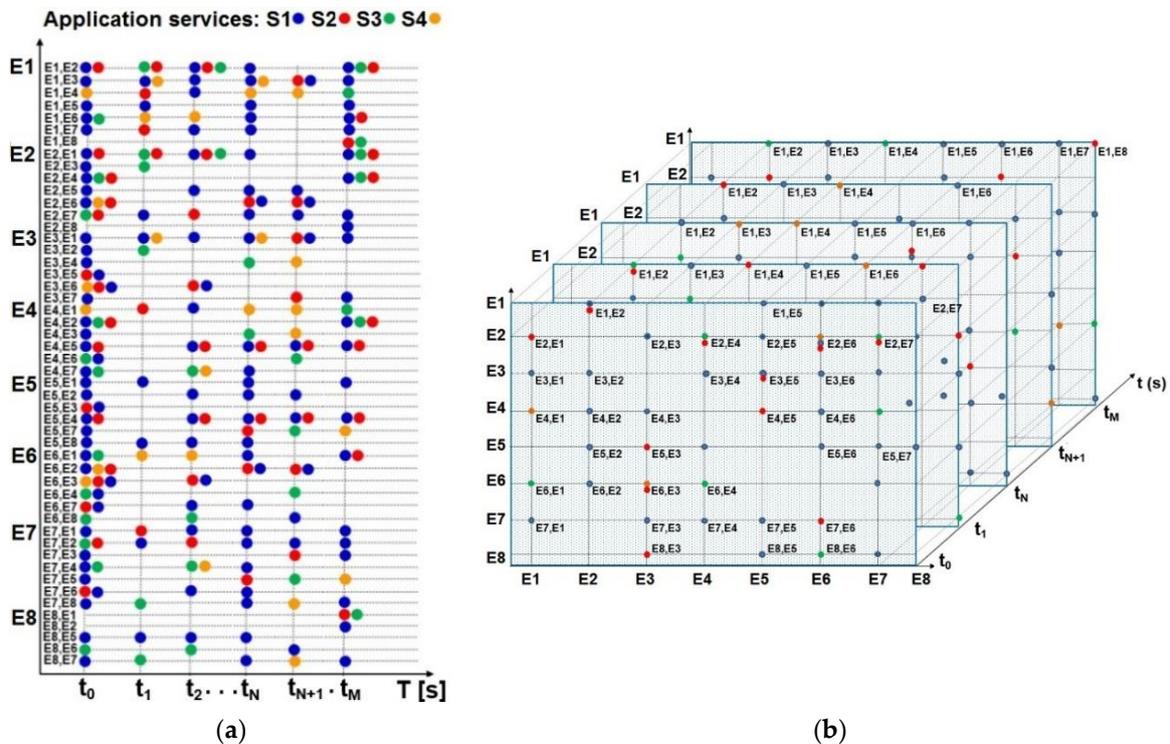
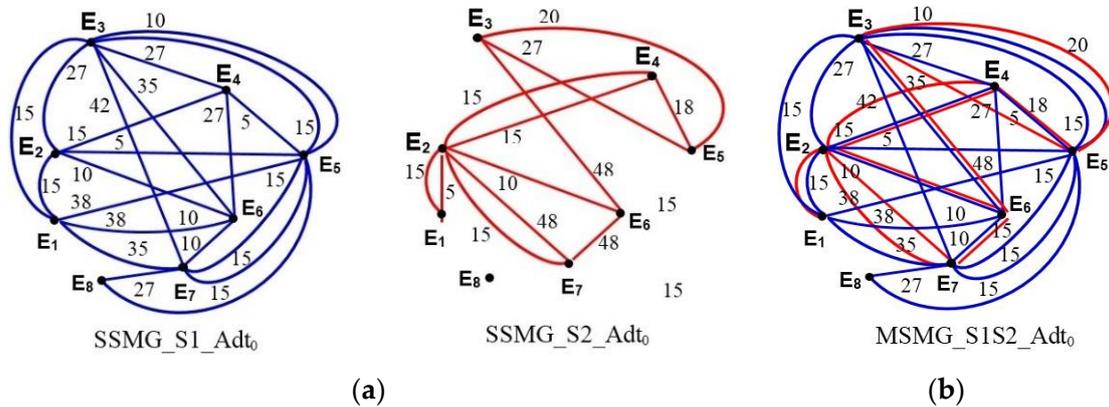


Figure 5. Activation and repetition of the application services S1 to S4 and the communication interactions among network elements E1 to E8: (a) timeline; (b) time plane.

#### 4.2. Multigraphs of Data Distribution in ITCN Network Traffic

The data exchanged by the applicable service Srv at each moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  are shown by presenting the network traffic as a multigraph (Figure 6a). The single-service multigraph (labeled SSMG\_Srv\_Ad $t$ ) shows the amount of data Adt (kbps or Mbps) exchanged between the network elements  $E_i$  and  $E_j$  ( $i \neq j$ ) by the application service Srv at time t. A single edge between the nodes (simple graphs)  $E_i$  and  $E_j$  ( $i \neq j$ ) represents the communication interaction between these network elements, where the amount of data Adt are distributed through the application service Srv at time t. The creation of all the single-service multigraphs between the network elements  $E_i$  and  $E_j$  ( $i \neq j$ ) through the application service Srv for each moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  enables the presentation of data exchanged during the commu-

nication time interval  $\Delta T = [t_0, t_m]$ . The total data exchanged between the nodes  $E_i$  and  $E_j$  ( $i \neq j$ ) through all application services, are  $Srv = (S1, S2, \dots, S_n)$  with time  $t$  representing the unification of all the single-service multigraphs formed previously into one multi-service multigraph (labeled  $MSMG\_S1S_n\_Adt$ ), as shown in Figure 6b. The multi-service multigraph enables the definition of network traffic among the ITCN's network elements at the observed moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  and enables the application of graph sampling theory to perform predictions, as in [22].



**Figure 6.** Data exchange multigraph among network elements  $E_1$  to  $E_8$  with the application services  $S_1$  (blue line) to  $S_2$  (red line) at time  $t_0$ : (a) single-service multigraph; (b) multi-service multigraph.

The creation of multi-service multigraphs for each moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  enables the presentation of the exchange of all data through all application services during the communication time interval  $\Delta T = [t_0, t_m]$ . A set of multi-service multigraphs allows one to define the total network traffic among the ITCN's network elements during the communication time interval  $\Delta T = [t_0, t_m]$ .

4.3. Matrix Associated with the ITCN Network Traffic Distribution Multigraph

The multigraph data distribution in network traffic of the ITCN is mathematically represented by the symmetric matrix  $T_{SSMG\_Srv\_Adt}$  in Equation (1) with integer terms and a diagonal of zero, where  $n$  is the number of network elements  $E_i$ . The associated symmetric matrix is formed by using a timeline or a time plane of the communication interactions (Figure 5) or by using a single-service multigraph (Figure 6a), such that

$$T_{SSMG\_Srv\_Adt} = \begin{bmatrix} 0 & Ad_{12}t & Ad_{13}t & Ad_{14}t & Ad_{15}t & \dots & Ad_{1n}t \\ Ad_{21}t & 0 & Ad_{23}t & Ad_{24}t & Ad_{25}t & \dots & Ad_{2n}t \\ Ad_{31}t & Ad_{32}t & 0 & Ad_{34}t & Ad_{35}t & \dots & Ad_{3n}t \\ Ad_{41}t & Ad_{42}t & Ad_{43}t & 0 & Ad_{45}t & \dots & Ad_{4n}t \\ Ad_{51}t & Ad_{52}t & Ad_{53}t & Ad_{54}t & 0 & \dots & Ad_{5n}t \\ Ad_{61}t & Ad_{62}t & Ad_{63}t & Ad_{64}t & Ad_{65}t & \dots & Ad_{6n}t \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \vdots \\ Ad_{n1}t & Ad_{n2}t & Ad_{n3}t & Ad_{n4}t & Ad_{n5}t & \dots & 0 \end{bmatrix} \quad (1)$$

where  $Ad_{ij}t$  is the amount of data distributed in the communication interactions between the nodes  $E_i$  and  $E_j$  ( $i \neq j$ ) with the application service  $Srv = (S_1, S_2, S_n)$  at the moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ . Figure 7 shows the single-service multigraph for data exchanged among the network elements  $E_1$  to  $E_8$  with the application service  $S_1$  at the moment of time  $t_0$ , and its associated symmetric  $8 \times 8$  matrix.

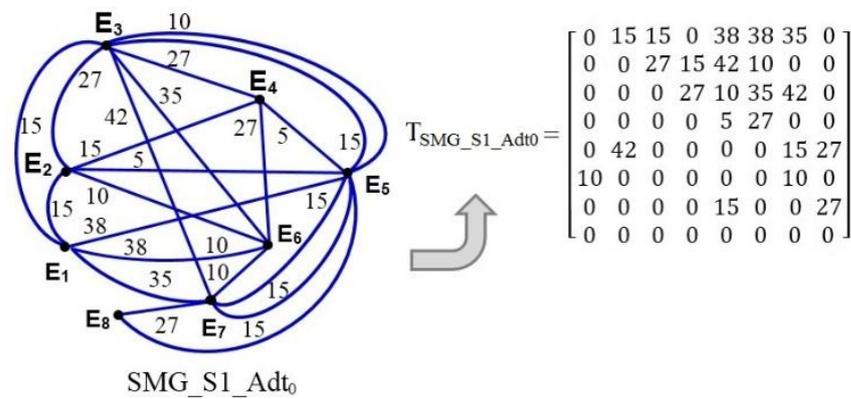


Figure 7. Data exchange single-service multigraph and its symmetric  $8 \times 8$  matrix.

For all single-service multigraphs, the set of associated matrices  $T_{SSMG\_Srv\_Adt}$  defines the matrix at each moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ , as in Equations (2) and (3):

$$T_{SSMG\_S1\_Adt_0} = \begin{bmatrix} 0 & Ad_{12}t_0 & Ad_{13}t_0 & Ad_{14}t_0 & Ad_{15}t_0 & \dots & Ad_{1n}t_0 \\ Ad_{21}t_0 & 0 & Ad_{23}t_0 & Ad_{24}t_0 & Ad_{25}t_0 & \dots & Ad_{2n}t_0 \\ Ad_{31}t_0 & Ad_{32}t_0 & 0 & Ad_{34}t_0 & Ad_{35}t_0 & \dots & Ad_{3n}t_0 \\ Ad_{41}t_0 & Ad_{42}t_0 & Ad_{43}t_0 & 0 & Ad_{45}t_0 & \dots & Ad_{4n}t_0 \\ Ad_{51}t_0 & Ad_{52}t_0 & Ad_{53}t_0 & Ad_{54}t_0 & 0 & \dots & Ad_{5n}t_0 \\ Ad_{61}t_0 & Ad_{62}t_0 & Ad_{63}t_0 & Ad_{64}t_0 & Ad_{65}t_0 & \dots & Ad_{6n}t_0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \vdots \\ Ad_{i1}t_0 & Ad_{i2}t_0 & Ad_{i3}t_0 & Ad_{i4}t_0 & Ad_{i5}t_0 & \dots & 0 \end{bmatrix} \quad (2)$$

$$T_{SSMG\_S1\_Adt_m} = \begin{bmatrix} 0 & Ad_{12}t_m & Ad_{13}t_m & Ad_{14}t_m & Ad_{15}t_m & \dots & Ad_{1n}t_m \\ Ad_{21}t_m & 0 & Ad_{23}t_m & Ad_{24}t_m & Ad_{25}t_m & \dots & Ad_{2n}t_m \\ Ad_{31}t_m & Ad_{32}t_m & 0 & Ad_{34}t_m & Ad_{35}t_m & \dots & Ad_{3n}t_m \\ Ad_{41}t_m & Ad_{42}t_m & Ad_{43}t_m & 0 & Ad_{45}t_m & \dots & Ad_{4n}t_m \\ Ad_{51}t_m & Ad_{52}t_m & Ad_{53}t_m & Ad_{54}t_m & 0 & \dots & Ad_{5n}t_m \\ Ad_{61}t_m & Ad_{62}t_m & Ad_{63}t_m & Ad_{64}t_m & Ad_{65}t_m & \dots & Ad_{6n}t_m \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \vdots \\ Ad_{i1}t & Ad_{i2}t & Ad_{i3}t & Ad_{i4}t & Ad_{i5}t & \dots & 0 \end{bmatrix} \quad (3)$$

The set of associated matrices  $T_{SSMG\_Srv\_Adt}$  enables one to define the function for the distribution of data in the network through the service  $Srv = (S1, S2, \dots, Sn)$  in the communication time interval  $\Delta T = [t_0, t_m]$ .

The variation in the value of the amount of data  $VarAdt$  distributed by the application service  $Srv = (S1, S2, \dots, Sn)$  during the communication time interval  $\Delta T = [t_0, t_m]$  is defined by the minimum and maximum values of the amount of data distributed among the network elements of the ITCN:

$$VarAdt = [Ad_{min}, Ad_{max}] \quad (4)$$

$$Ad_{min} = \min\{Ad_{ij}t_0, \dots, Ad_{ij}t_m\}, i = [1, n] \quad j = [1, n] \quad i \neq j \quad (5)$$

$$Ad_{max} = \max\{Ad_{ij}t_0, \dots, Ad_{ij}t_m\}, i = [1, n] \quad j = [1, n] \quad i \neq j. \quad (6)$$

For the multi-service multigraph, the associated symmetric  $n \times n$  matrix  $T_{MSMG\_S1Sn\_Adt}$  of data distribution shown in Equation (7) is formed. The value of the distribution function of the total amount of data distributed through all application services  $Srv$  at the

moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  is defined by the associated symmetric matrix  $T_{\text{MSMG\_S1Sn\_Ad}t}$ .

$$T_{\text{MSMG\_S1Sn\_Ad}t} = \begin{bmatrix} 0 & sAd_{12}t & sAd_{13}t & sAd_{14}t & \dots & sAd_{1n}t \\ sAd_{21}t & 0 & sAd_{23}t & sAd_{24}t & \dots & sAd_{2n}t \\ sAd_{31}t & sAd_{32}t & 0 & sAd_{34}t & \dots & sAd_{3n}t \\ sAd_{41}t & sAd_{42}t & sAd_{43}t & 0 & \dots & sAd_{4n}t \\ sAd_{51}t & sAd_{52}t & sAd_{53}t & sAd_{54}t & \dots & sAd_{5n}t \\ sAd_{61}t & sAd_{62}t & sAd_{63}t & sAd_{64}t & \dots & sAd_{6n}t \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots \\ sAd_{n1}t & sAd_{n2}t & sAd_{n3}t & sAd_{n4}t & \dots & 0 \end{bmatrix} \tag{7}$$

where  $sAd_{ij}t$  is the total amount of data distributed in communication interactions between nodes  $E_i$  and  $E_j$  ( $i \neq j$ ) with all the activated application services  $Srv = (S_1, S_2, \dots, S_n)$  at the moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ , where:

$$sAd_{ij}t = \sum_{S_1}^{S_n} \sum_{i=1}^n \sum_{j=1}^n Ad_{ij}t, \quad i \neq j, \quad t = (t_0, t_1 \dots t_{N-1}, t_N, t_m) \tag{8}$$

The set of associated matrices  $T_{\text{MSMG\_S1Sn\_Ad}t}$  at each moment of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  defines a set of multi-service multigraphs. The set of associated symmetric matrices  $T_{\text{MSMG\_S1Sn\_Ad}t}$  enables the definition of the value of the distribution function of the total amount of data distributed through all the application services  $Srv$  during the communication time interval  $\Delta T = [t_0, t_m]$ , such that

$$T_{\text{MSMG\_S1Sn\_Ad}t_0} = \begin{bmatrix} 0 & sAd_{12}t_0 & sAd_{13}t_0 & sAd_{14}t_0 & \dots & sAd_{1n}t_0 \\ sAd_{21}t_0 & 0 & sAd_{23}t_0 & sAd_{24}t_0 & \dots & sAd_{2n}t_0 \\ sAd_{31}t_0 & sAd_{32}t_0 & 0 & sAd_{34}t_0 & \dots & sAd_{3n}t_0 \\ sAd_{41}t_0 & sAd_{42}t_0 & sAd_{43}t_0 & 0 & \dots & sAd_{4n}t_0 \\ sAd_{51}t_0 & sAd_{52}t_0 & sAd_{53}t_0 & sAd_{54}t_0 & \dots & sAd_{5n}t_0 \\ sAd_{61}t_0 & sAd_{62}t_0 & sAd_{63}t_0 & sAd_{64}t_0 & \dots & sAd_{6n}t_0 \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots \\ sAd_{n1}t_0 & sAd_{n2}t_0 & sAd_{n3}t_0 & sAd_{n4}t_0 & \dots & 0 \end{bmatrix} \tag{9}$$

$$T_{\text{MSMG\_S1Sn\_Ad}t_m} = \begin{bmatrix} 0 & sAd_{12}t_m & sAd_{13}t_m & sAd_{14}t_m & \dots & sAd_{1n}t_m \\ sAd_{21}t_m & 0 & sAd_{23}t_m & sAd_{24}t_m & \dots & sAd_{2n}t_m \\ sAd_{31}t_m & sAd_{32}t_m & 0 & sAd_{34}t_m & \dots & sAd_{3n}t_m \\ sAd_{41}t_m & sAd_{42}t_m & sAd_{43}t_m & 0 & \dots & sAd_{4n}t_m \\ sAd_{51}t_m & sAd_{52}t_m & sAd_{53}t_m & sAd_{54}t_m & \dots & sAd_{5n}t_m \\ sAd_{61}t_m & sAd_{62}t_m & sAd_{63}t_m & sAd_{64}t_m & \dots & sAd_{6n}t_m \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots \\ sAd_{n1}t_m & sAd_{n2}t_m & sAd_{n3}t_m & sAd_{n4}t_m & \dots & 0 \end{bmatrix} \tag{10}$$

The elements for defining the data distribution function are realized by forming all the sets of associated single-service matrices  $T_{\text{SSMG\_Srv\_Ad}t}$  and all the sets of the associated multi-service matrices  $T_{\text{MSMG\_S1Sn\_Ad}t}$ .

### 5. Generating the Data Distribution Function in the ITCN by Sampling Multigraphs

The function distribution of the amount of data in the time interval (pdf( $\Delta T$ )) for implementation in the OPNET simulation model is defined by sampling the single-service and multi-service multigraphs of data distribution. Sampling the multigraphs is equivalent to sampling the associated symmetric matrices given in [23]. The associated symmetric matrix is sampled using the sequential importance sampling (SIS) method for sampling multigraphs given in [20]. The value of the estimated distribution function represents the

content of the multigraph and is determined by applying the asymptotic approximation given in [20,21].

Additionally, the distribution function for the approximation of multigraphs is defined by using graphs and weighting coefficients and applying the methods given in [20,24].

The matrix  $T_{SSMG}(t) = T_{SSMG\_Srv\_Adt}$  belongs to the set of associated symmetric matrices related to the distribution of data between the network elements  $E_i$  and  $E_j$  ( $i \neq j$ ) with the application service  $Srv$  at the moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ , where  $\Sigma T$  is the number of matrices in the set. The distribution function  $q(T_{SSMG}(t)) > 0$  for the matrix  $T_{SSMG}(t)$  defines the amount of data for distribution between the network elements  $E_i$  and  $E_j$  ( $i \neq j$ ) through the application service  $Srv$  at the moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$ . The estimated value of the distribution function is:

$$E_q \left[ \frac{1}{q(T_{SSMG}(t))} \right] = \sum_T \frac{1}{q(T_{SSMG}(t))} q(T_{SSMG}(t)) = |\Sigma T| \tag{11}$$

$$|\Sigma T| = \frac{1}{n} \sum_{i=1}^n \frac{1}{q(T_{SSMG}(t))} \tag{12}$$

The distribution function  $q(T_{SSMG}(t))$  is determined with the test distribution function  $q(\cdot)$  by sampling the  $T_{SSMG}(t)$  matrix column by column  $(c_1, c_2, \dots, c_n)$ , using the method and procedure in [20] and [21]. Here,  $q(T_{SSMG}(t))$  is represented by:

$$q(T_{SSMG}(t) = (c_1, c_2, \dots, c_n)) = q(c_1)q(c_1|c_2) \dots q(c_n|c_{n-1} \dots, c_1)_{t_1, \dots, t_{N-1}} \tag{13}$$

The sum of the row margin ( $d_i$ ) of the matrix ( $n \times n$ ), denoted  $d^{(2)}, d^{(3)}, \dots, d^{(i)}$ , and the updated row margins of the  $(n-1) \times (n-1)$  submatrix are determined for the matrix  $T_{Ei}(t)$ .

The procedure of sampling and removing the matrix columns in  $T_{SSMG}(t)$  is repeated until all the columns  $(c_1, c_2, \dots, c_n)$  have been sampled. The value of each margin of the row ( $d_i$ ) and the total margin of the matrix ( $M$ ) is calculated according to the following:

$$d_i = \sum_{j=1}^n \alpha_{ij}, \quad i = [1, n] \tag{14}$$

$$d = (d_1, d_2, d_3, \dots, d_n) \tag{15}$$

$$d^{(2)} = (d_2 - \alpha_{21}, d_4 - \alpha_{42}, \dots, d_n - \alpha_{n2}) \tag{16}$$

$$d^{(i)} = (d_i - \alpha_{i,i-1}, d_{i+1} - \alpha_{i+1,i-1}, \dots, d_n - \alpha_{n,i-1}), \quad i = [2, n] \tag{17}$$

$$M = \sum_{i=1}^n d_i \tag{18}$$

For the  $T_{SSMG}(t)$  matrix ( $t$ ), the number of multigraphs  $|\Sigma d|$  is calculated. Submatrices are formed by removing columns. For forming the submatrices, the number of multigraphs  $|\Sigma d^{(i)}|$  is calculated, which corresponds to the associated submatrix. Based on the asymptotic approximation given in [20] and [21], the expression for  $|\Sigma d|$  and for  $|\Sigma d^{(i)}|$  is performed.

$$|\Sigma d| \sim \Delta_d \equiv \frac{f(M)}{\prod_{i=1}^n d_i!} e^{a(d)} \tag{19}$$

$$f(M) = M! / \left[ \left( \frac{M}{2} \right)! 2^{\frac{M}{2}} \right] \tag{20}$$

$$a(d) = \left( \sum_i \binom{d_i}{2} / M \right)^2 - \sum_i \binom{d_i}{2} / M \tag{21}$$

$$|\Sigma d^{(2)}| \sim \Delta_{d^{(2)}} \equiv \frac{f(M - 2d_1)}{\prod_{i=1}^n (d_i - \alpha_{i1})!} e^{a(d^{(2)})} \tag{22}$$

The expression obtained for each column  $(c_1, c_2, \dots, c_n)$  of the  $T_{SSMG}(t)$  matrix determines the marginal distribution function of each column  $p(c_i) \sim q(c_i)$ . The marginal distribution function represents the derived distribution function  $q(T_{SSMG}(t))$ .

$$p(c_1 = (0, \alpha_{21}, \dots, \alpha_{n1})) = \frac{|\sum d^{(2)}|}{|\sum d|} \tag{23}$$

$$p(c_2) = \frac{|\sum d^{(3)}|}{|\sum d|} \tag{24}$$

$$p(c_{n-1}) = \frac{|\sum d^{(n)}|}{|\sum d|} \tag{25}$$

Combining the Expressions in (19) and (22) derives an expression for  $q(c_1)$ :

$$q(c_1 = (0, \alpha_{21}, \dots, \alpha_{n1})) = \frac{1}{\prod_{i=1}^n (d_i - \alpha_{i1})!} e^{a(d^{(2)})} \tag{26}$$

The expressions for  $q(c_1 | c_2), \dots, q(c_n | c_{n-1}, \dots, c_1)$  are derived in the same way. The value of  $q(T_{SSMG}(t))$  is calculated from the obtained values. The procedure given in [19] evaluates the sampling efficiency of the matrix and the accuracy of the derived distribution function  $q(T_{SSMG}(t))$  in relation to the marginal distribution  $p(T_{SSMG}(t))$ . The value of the standard estimation error  $\mu$  and the difference between the obtained values of  $cv^2$  is used to calculate the following expression:

$$\hat{\mu} = \frac{\sum_{i=1}^n f(T_{SSMGi}(t)) \frac{p(T_{SSMG}(t))}{q(T_{SSMG}(t))}}{\sum_{i=1}^n \frac{p(T_{SSMG}(t))}{q(T_{SSMG}(t))}} = \frac{\sum_{i=1}^n f(T_{SSMG}(t)) \frac{\frac{1}{q(T_{SSMG}(t))}}{\frac{1}{q(T_{SSMG}(t))}}}{\sum_{i=1}^N \frac{1}{\frac{1}{q(T_{SSMG}(t))}}} = \frac{\sum_{i=1}^N f(T_{SSMG}(t)) \frac{1}{q(T_{SSMG}(t))}}{\sum_{i=1}^N \frac{1}{q(T_{SSMG}(t))}} \tag{27}$$

The use of weight coefficients (weights)  $\omega_i$  calculated by the procedure given in [20] and [22] realizes the correction and adjustment of values between the derived distribution function  $q(T_{SSMG}(t))$  and the marginal distribution  $p(T_{SSMG}(t))$ .

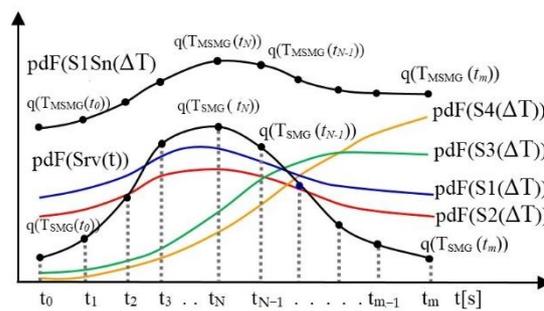
The application of the previous procedure to all single-service multigraphs and their associated matrices defines the change in the amount of data for distribution in the network. The change in the amount of distributed data is realized between network elements  $E_i$  and  $E_j$  ( $i \neq j$ ) through the application service  $Srv$  in the communication time interval  $\Delta T = [t_0, t_m]$ . The calculated values of the distribution function form a set of values of the distribution function  $q(T_{SSMG}(t))$ . These values enable one to define the data distribution function  $pdf(Srv(t))$  of the application service  $Srv$  in the communication time interval  $\Delta T = [t_0, t_m]$ :

$$\{q(T_{SSMG}(t))\} \Rightarrow pdf(q(T_{SSMG}(t))), \mathbf{t} = (t_0, t_1, \dots, t_{N-1}, t_N, t_m) \tag{28}$$

$$pdf(Srv(\Delta T)) = pdf(q(T_{SSMG}(t))), Srv = (S_1, \dots, S_n) . \tag{29}$$

The same procedure applies to multi-service multigraphs. The calculated values of the distribution function  $pdf(q(T_{MSMG}(t)))$  define the data distribution function  $pdf(S_1 S_n(\Delta T))$  of all application services  $Srv$  in the communication time interval  $\Delta T = [t_0, t_m]$ . The calculated values form a set of values  $\{q(T_{SSMG}(t))\}$ . The use of values from the set of values  $\{q(T_{SSMG}(t))\}$  thus formed enables the creation of graphs of the distribution function  $pdf(Srv(\Delta T))$ . Graphically, the values are connected in the order of the moments of time  $t = (t_0, t_1, \dots, t_{N-1}, t_N, t_m)$  to which the values refer.

The graph of the data distribution function  $pdf$  (Figure 8) shows the regularity of the time of the change in the amount of data for distribution among the ITCN's network elements through the application service  $Srv$  in the communication time interval  $\Delta T = [t_0, t_m]$ .



**Figure 8.** The graph of the data distribution function pdf through the application service Srv in the communication time interval  $\Delta T = [t_0, t_m]$ .

The same procedure is used to define the graph of the data distribution function of all the application services  $\text{pdf}(S1Sn(\Delta T))$  in the communication time interval  $\Delta T = [t_0, t_m]$ . Determining the similarity of the graphs of the function  $\text{pdf}(Srv(t))$  to the graphs of the known distribution functions (exponential, Poisson, Normal (Gaussian), uniform, Weibull, etc.) given in [8,10,16,17] allows one to identify the derived distribution function  $\text{pdf}(Srv(t))$ . The identification of  $\text{pdf}(Srv(t))$  as a known distribution function enables the selection of the existing distribution function in the OPNET simulation model and the application of the parameter values from the graphs (Figure 8). If no similarity is found, the use of software tools integrated into the OPNET software allows one to import graphics of the derived distribution function  $\text{pdf}(Srv(t))$ . This way, the distribution function can be used as a newly defined distribution for realization of the ITCN simulation model.

## 6. Conclusions and Further Research

The application of multigraphs for describing data distribution in an ITCN was described in this study. It enabled a more accurate definition of communication events in the network and a mathematical description of the network traffic. The described method of applying multigraphs is primarily intended for the development of a simulation model of networks with deterministically defined and controlled communication. The method of applying multigraphs is also possible in networks with stochastically generated network traffic, which requires a definition of the variations in the amount of data to be generated between the source and destination. Achieving more accurate results of simulating the predicted communication in an ITCN enables the integration of all the derived distribution functions for a description of the network traffic into the OPNET simulation model. Based on the analysis of the results of realized discrete OPNET simulations of an ITCN, we propose the application of single-service multigraphs, and the derivation and use of the distribution functions of  $\text{pdf}(Srv(t))$  of data distribution for each application service. For realizing the simulation of network traffic flows, we propose the use of multi-service multigraphs, and the creation and use of data distribution functions  $\text{pdf}(S1Sn(\Delta T))$  of all services at the same time. In future research, we will analyze the correlation of the data distribution functions with functions that describe individual network parameters (connectivity, capacity, etc.). This research will enable predictions in the design and optimization of an ITCN using a simulation model.

**Author Contributions:** Methodology and algorithm development: S.M. and V.M.; simulation, data analysis, testing and validation: S.M. and I.P.; Multigraph analysis: K.P.-P. and G.R.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Miletic, S.; Milosevic, M.; Mladenovic, V. A New Methodology for Designing of Tactical Integrated Telecommunication and Computer Networks for OPNET Simulation. In Proceedings of the 9th International Scientific Conference on Defensive Technologies, OTEH 2020, Belgrade, Serbia, 15–16 October 2020; Technical Review. Volume 70, pp. 35–40.
2. Miletic, S.; Mladenovic, V.; Pokrajac, I. Application of multigraph sampling method in network traffic design of simulation model of Integrated Telecommunication and Computer Network. *E3S Web Conf.* **2021**, *279*, 02011. [[CrossRef](#)]
3. Tatarnikova, T.; Sikarev, I.; Karetnikov, V.; Butsanets, A. Statistical research and modeling network traffic. *E3S Web Conf.* **2021**, *244*, 07002. [[CrossRef](#)]
4. Antoniuo, I.; Ivanov, V.V.; Zrellov, P.V. Statistical model of network traffic. *Phys. Part. Nucl.* **2004**, *35*, 530. [[CrossRef](#)]
5. Chen, T.M. Chapter in The Handbook of Computer Networks. In *Network Traffic Modeling*; Hossein, B., Ed.; Wiley: Hoboken, NJ, USA, 2007.
6. Dymora, P.; Mazurek, M.; Strzalka, D. Computer network traffic analysis with the use of statistical self-similarity factor. *Ann. UMCS Inform. AI XIII* **2013**, *2*, 69–81. [[CrossRef](#)]
7. Alsamar, M.; Parisi, G.; Clegg, R.; Zakhleniuk, N. On the distribution of traffic volumes in the Internet and its implications. *arXiv* **2019**, arXiv:1902.03853v1 [cs.NI].
8. Leemis, L. *Input Modeling Techniques for Discrete-Event Simulations*; Department of Mathematics, The College of William & Mary: Williamsburg, VA, USA, 2021; pp. 23187–28795.
9. Sanchez, P.J. Fundamentals of simulation modeling. In Proceedings of the Winter Simulation Conference, Washington, DC, USA, 9–12 December 2007.
10. Chandrasekaran, B. Survy of Network Traffic Models. Available online: [https://www.cse.wustl.edu/~jain/cse567-06/ftp/traffic\\_models3/index.html](https://www.cse.wustl.edu/~jain/cse567-06/ftp/traffic_models3/index.html) (accessed on 1 May 2021).
11. Markelov, O.; Duc, V.N.; Bogachev, M. *Statistical Modeling of the Internet Traffic Dynamics: To Which Extent Do We Need Long-Term Correlations*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 485, pp. 48–60. [[CrossRef](#)]
12. Malyeyeva, O.; Davydovskiy, Y.; Kosenko, V. Statistical Analysis of Data on the Traffic Intensity of Internet Networks for the Different Periods of Time. In Proceedings of the Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019), Zaporizhzhia, Ukraine, 15–19 April 2019; pp. 897–910.
13. Davydovskiy, Y.; Reva, O.; Artiukh, O.; Kosenko, V. Simulation of Computer Network Load Parameters over a Given Period of Time. In *Innovative Technologies and Scientific Solutions for Industries*; Quarterly Scientific Journal: Kharkiv, Ukraine, 2019; ISSN 2524-2296. [[CrossRef](#)]
14. Barner, K.; Gonzalo, R.A. *Processing Theory, Methods and Applications*, CRC Press LLC, 2000 N.W.; Corporate Blvd.: Boca Raton, FL, USA, 2004; ISBN 0-8493-1427-5.
15. Arce, G.R. *Nonlinear Signal Processing: A Statistical Approach*; Wiley and Sons: Hoboken, NJ, USA, 2004.
16. Schmidt, R.; De, O.; Sadre, R.; Pras, A. Gaussian traffic revisited. In *2013 IFIP Networking Conference*; IEEE: Piscataway, NJ, USA, 2013; ISBN 978-3-901882-55-5.
17. Manaseer, S.; Al-Nahar, O.M.; Hyassat, A.S. Network traffic modeling. *Int. J. Recent Technol. (IJRTE)* **2019**, *7*.
18. Gongx, Y.; Wang, X.; Malboubi, M. Towards Accurate Online Traffic Matrix Estimation in Software-Defined Networks. In Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Network Research, Santa Clara, CA, USA, 17–18 June 2015; pp. 1–7.
19. Mukhin, V.; Romanenkov, Y.; Bilokin, J.; Rohovyi, A.; Kharazii, A.; Kosenko, V.; Kosenko, N.; Su, J. The Method of Variant Synthesis of Information and Communication Network Structures on the Basis of the Graph and Set-Theoretical Models. *Int. J. Intell. Syst. Appl.* **2017**, *11*, 42–51. [[CrossRef](#)]
20. Eisinger, R.D.; Chen, Y. Sampling strategies for conditional inference on multigraphs. *Stat. Its Interface* **2018**, *11*, 649–656. [[CrossRef](#)]
21. Chen, Y.; Diaconis, P.; Holmes, S.P.; Liu, J.S. Sequential Monte Carlo methods for statistical analysis of tables. *J. Am. Stat. Assoc.* **2005**, *100*, 109–120. [[CrossRef](#)]
22. Sardellitti, S.; Barbarossa, S.; Di Lorenzo, P. *Enabling Prediction via Multi-Layer Graph Inference and Sampling*, Auckland University of Technology; IEEE: Piscataway, NJ, USA, 2020; pp. 1–4.
23. Lau, D.L.; Arce, G.R.; Parada-Mayorga, A.; Dapena, D.; Pena-Pena, K. Blue-Noise Sampling of Graph and Multigraph Signals: Dithering on Non-Euclidean Domains. *IEEE Signal Processing Mag.* **2020**, *37*, 31–42. [[CrossRef](#)]
24. Barrat, A.; Barthelemy, M.; Satorras, R.P.; Vespignani, A. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3747. [[CrossRef](#)] [[PubMed](#)]