Mathematical Problems of Computer Science 48, 98-104, 2017.

Consideration of Congestion Situations in Telecommunication Networks and Methods of Their Processing as a Special Type of Emergency Situations

Hamlet H. Harutyunyan and Hmayak A. Avanesyan

Armenian State Pedagogical University after Kh. Abovyan
National Polytechnic University of Armenia
e-mail: hamhar@lycos.com, hmayak.avanesyan@gmail.com

Abstract

This article discusses the methods of struggle with congestion situations in telecommunication networks, considering them as a special type of emergency situations, which frequently occur in computing systems.

Keywords: Telecommunication network, Congestion situation, emergency, Network traffic, Intensity of traffic, Data transmission.

1. Introduction

The problem of struggle with congestion situations is one of the most difficult and important problems of telecommunication networks. Many scientific works have been devoted to this problem. Currently, the main directions of research are conducted generally in the following areas:

- increase the performance of switching centers (nodes);
- increase the bandwidth of communication channels;
- improvement of data transfer protocols;
- creation of effective traffic processing algorithms;
- development of methods for preventing, detecting, processing and eliminating of congestion situations with minimal data loss.

In this article, the congestion situation is considered as an emergency in switching centers. In accordance with this idea, the methods of struggle with congestion situations are considered as a special case in general methodology of emergency situations processing. Such approach to congestion situations extends the methodology of their research and processing, using the existing methodologies of emergency management.

2. Congestions as a Special Type of Emergency Situations

The concept of congestion situation is defined in telecommunication networks literature in different ways [1,2,4]. Mostly it is associated with network traffic increase.

Here is a brief description of the process of occurrence of congestion situation. The congestion occurs when the current load of system exceeds its maximum load capacity. While the current system load is below the acceptable value, the intensity of the incoming traffic is equal or almost equal to the intensity of the outgoing traffic in the switching center. As soon as the current load approaches and exceeds the system load capacity, the intensity of the incoming traffic becomes more than the intensity of the outgoing traffic. A queue of packets waiting to be processed accumulates in the switching center and the delay of their processing begins to increase. The more the load increases, the longer the queue becomes and, hence, the longer the delay period becomes. With the continuation of this process, the queue length grows to the limit, after that, new incoming packets are not queued but discarded. Without having received a confirmation from the recipient, the sender starts to resend the packets, which in turn, further increases the load of the receiving center. In this situation, the bandwidth capability of the switching center begins to fall and the outgoing traffic decreases sharply. In the end, it tends to zero, which brings about the switching center dysfunction. This situation in the scientific literature is called congestion [1,5].

The negative consequences of congestion situations are:

- telecommunication network characteristics deterioration;
- transmission of data delays and losses;
- bandwidth reduction of switching centers;
- partial or complete dysfunction of the switching center.

In this context, the state of congestion resembles the classical definition of the computing systems emergency state, when due to hardware or software errors, data losses and partial or complete system faults occur [1]. Despite the similarities, different types of emergencies are different by nature, and there is no universal method for their detection and processing. Therefore, for every type of emergency, special methods are required to develop appropriate for their nature, conditions of occurrence and impact on the functioning of the system.

The general emergency management strategy involves the following procedures:

- prevention of occurrence of emergency situations;
- control and detection of emergency situations;
- processing of detected emergency situations and restoring the normal functioning mode of the system;
- elimination of consequences of emergency conditions.

The strategy of struggle with congestions also includes the same procedures mentioned above. However, despite the general methodology, congestion as an emergency situation has a completely different nature. Unlike the classic concept of emergency situations, congestion can occur in a state of full operational capability of hardware and software systems and complete correctness of information. The congestion situation occurs when the rate of incoming traffic exceeds a certain threshold value corresponding to the switching center maximum value throughput.

The congestion situation can be eliminated itself, when the incoming traffic becomes less than the specified threshold value. In that case the congestion situation is like a “transient fault” in hardware.

The concept of congestion is closely related to the data processing in real time mode, when there is a limitation for packets transmission time, and violation of limits is considered as an error. The absence of time limits formally means that the tasks queue length can be unlimited and the system can cope with the processing of traffic of any intensity. In this case, the congestion situation does not occur, because there is no violation of the time limit for data processing. Of course, such statement of question is purely theoretical. In practice, the queue length cannot be infinite, and, besides, any service must be completed within a reasonable time corresponding to the service specificity. If the time is exceeded, the interest of users to this service decreases.

It is important to note that the congestion situation can occur not only with the increase in incoming traffic, but also with a decrease in the bandwidth capability of the switching center due to hardware and software failures. Those failures can be caused by system characteristics deterioration. In this case, the system bandwidth can be decreased and become insufficient for processing tasks in proper time. That leads to a congestion situation.

It should be noted that, due to failures in hardware and communication channels, the delays of processing packets in the queues sometimes can exceed the permissible value. However, that is not a signal of congestion. The signal of violation of the permissible delay time is a necessary but insufficient condition for fixing of congestion situation. Therefore, in such cases the danger of congestion occurs when the intensity of failures exceeds a preset threshold. Such a situation can arise due to the problems in the communication channels or congestion in the receiving centers.

The definition of congestion situation due to the conditions of processing tasks in real time is presented in [3].

3. The "sensitivity" of Switching Centers to the Congestion Situation

Telecommunication network has complex structures, consisting of many switching centers, which provide data transfer through the network. In packet switching mode switching centers have different "sensitivity" to the congestion situation. The probability of congestions depends on many factors, some of which are given below.

The efficiency of use of system and network resources. The switching centers as component parts of the common route usually have different loads. It depends on their bandwidth, number of routes, crossing in the switching center and intensity of the flow on them. The probability of congestion at any time depends on the difference between the bandwidth and the current load of the switching center (loading reserve).

$C_r = (C - C_t)$, where

C_r is the loading reserve, $C -$ the switching center bandwidth, C_t - the current load.

The less is this difference, the less free resources are left, and the greater is the probability of congestion. It is obvious, that in case of increasing traffic on the route, the congestion situation will occur in the switching center, which has less C_r.

The rank of the switching center. The switching centers depending on their position in a network topology, have different roles in the functioning of the network. The role of any switching center determines its "importance" in ensuring network bandwidth. In networks terminology the "importance" of switching nodes is characterized by the parameter "rank" of the node, which is determined by the number of communication channels connected to this node.

We can use the above-mentioned term "rank" for determination of switching centers "importance" in the topology of the telecommunication network. In this case, the parameter "rank" of the switching center can be determined by the number of possible routes passing through it.

The network topology can be represented as a graph. The graph vertices represent switching centers, and edges to communication channels. In the terminology of graph theory the parameter "rank" is the degree of the vertex of the graph. Obviously, under the same bandwidth capability switching centers, the probability of congestion will be more in a switching center with a higher rank. On the other hand, for equal values of the rank of switching centers, the probability of congestion is greater at the center with a lower bandwidth capability. An example of a graph of a network topology is shown in Fig. 2.1.

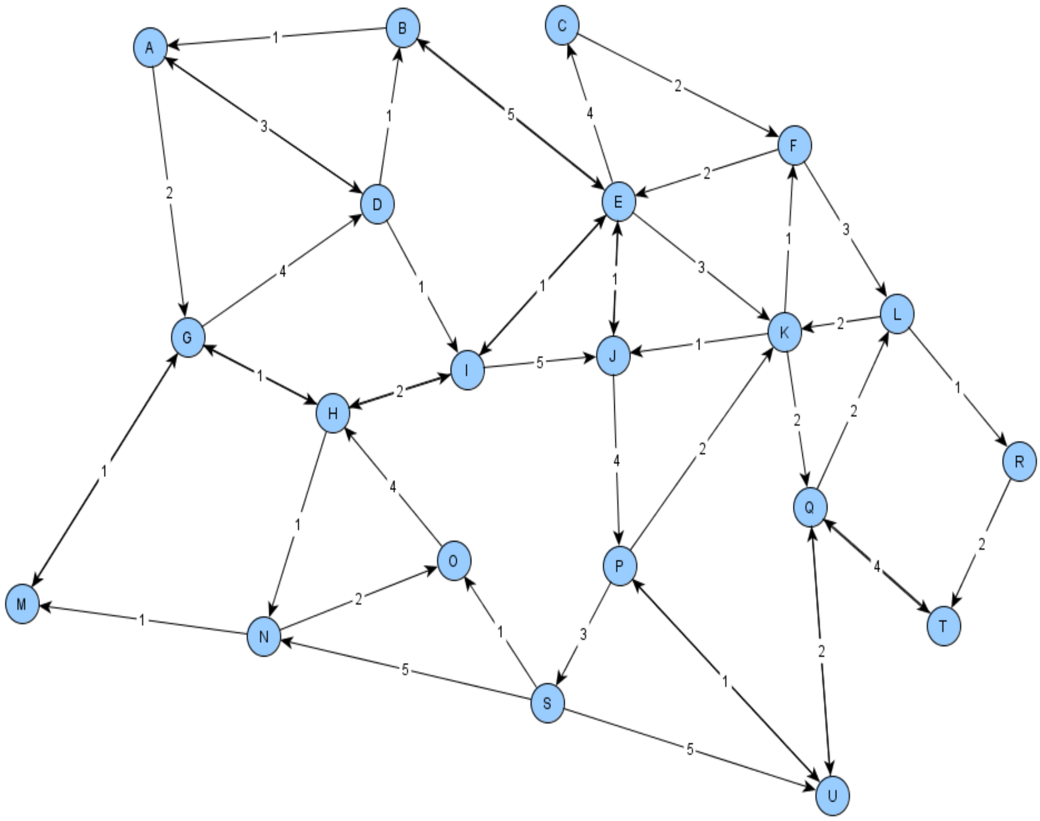

Fig. 2.1. Example of a graph topology of a telecommunication network.

In this example, switching centers have different "ranks": the rank of the switching center *C* is equal to 2, *K* is 6. It is obvious that the switching center number *K* is more significant in the network topology than the switching center number *C*.

The reliability of the switching center. The switching centers belong to the class of complex systems that are able to provide system functioning in case of fault of some components. The faults in hardware in such systems lead to the degradation of a system or functional characteristics. A fault in any component can lead to the state of partial or complete fault of system depending on the possibility of continuation of execution of system target functions. For switching center, a partial fault may lead to reduction of system and network resources, reducing the performance of the computer system or bandwidth of communication channels, the partial destruction of routes. In this connection, the switching center can get into a state of congestion even the intensity of incoming traffic is acceptable for a full configuration of system.

Sustainability of communication channels to errors. Increased intensity of failures in communication channels leads to increasing of number of received packets with errors and to increasing of number of resending of packets. In the end, it may lead to increased queue lengths in the buffer file and to the congestion of the sender center. The probability of overloading depends on the fault- tolerance of the channels.

4. The Influence of Switching Center Congestion Situation to the Quality of Functioning of the Entire Network

Congestion delays the flows on all routes passing through the switching center. It means that the local congestion in any switching center may cause congestion in other centers, which are connected to that center with common routs. This process can cover other switching centers and spread to levels of the subnet and the entire network.

The magnitude of the effect of congestion in any switching center on the performance of the whole network is determined by two factors:

1. The level of decreasing of network bandwidth.
2. The number of destroyed routes.

Congestion situation leads to an afunctional state of switching center for a while, which, in turn, leads to the change in the topology of network. To determine the impact of faults of a switching center on the network bandwidth, we introduce a "parameter" weight for any switching center.

We suggest the following method of calculating the parameter "weight" for switching centers.

1. To estimate the network bandwidth at the maximum allowed traffic over all possible routes *(C).*
2. To disconnect i-th switching center and to estimate the network bandwidth (C_i) again.

The "weight" of the i-th center is denoted by *Pi*.

$$P_i = 1 - \frac{C_i}{C},$$

P_i can vary in the range $0 \leq P_i \leq 1$;

If disabling the i-th switching center does not change the network bandwidth, then $P_i=0$ $(C_i=C)$. It can be in case, when there are reserves in the network topology and the output of the switching center from a working configuration does not affect the network bandwidth. If disabling i-th center *Pi=1 (Ci=0),* the network stops functioning, it means that a fault of i-th center leads to a complete fault of the whole network.

The more the *Pi* value, the greater the magnitude of the impact of i-th center on the functioning of the whole network.

The fault of the switching center can reduce the number of possible routes in the network, increase the load on other routes, change the network connectedness and stop transfer data by some routes.

When evaluating the impact of a switching center fault on the performance of the network, it should be taken into account not only the reducing number of possible routes, but the ‘importance” of the destroyed routes.

It is important to distinguish congestion due to lack of own resources and congestion due to receiving centers. Criteria for detection and a strategy of processing of both types of congestion situations are totally different.

The impact of congestion in a switching center to others also depends on the duration of congestion situation. We introduce the notion of "permissible duration of congestion situation”. This is a time interval, after which the sending center fixes the fact of congestion state in receiving center. If during that time the congestion situation is eliminated, the functioning mode will be restored.

The congestion duration depends on the behavior of the incoming traffic. When the incoming traffic intensity becomes lower than the system bandwidth, the overload is eliminated by itself. Taking into account the self-similar nature of network traffic, it can be assumed, that in such a short time as permissible duration interval, the probability of maintaining high intensity is higher than the probability of its decrease [6]. The permissible duration can be determined by the parameter of permissible delays of processing of packets in switching center, which for different traffic categories is different.

The strategy of processing of different types of emergency situations is discussed in scientific literature. However, the congestion situation is considered usually in the context of resources planning and flows management issues, but not within the framework of the emergency management strategy. All these works are mainly aimed at improving the efficiency of planning and managing the resources of telecommunication networks and also improving the data transmission procedures. It, of course, increases the efficiency of system and its stability to overload. However, with all this, the probability of congestion situation will still remain. Therefore, in addition to the conventional approach to congestion, it is advisable to consider the congestion as a special type of emergency situation in telecommunication networks and to apply the whole methodology of dealing with emergency situations.

References

[1] A. S. Tanenbaum, David J. Wetherall, Computer Networks, Fifth edition, Prentice Hall Press Upper Saddle River, NJ, USA 2010.

[2] I. Villy, T*eletraffic Engineering and Network Planning*, Technical University of Denmark, Ørsteds Plads, Denmark, 2015.

[3] H. Harutyunyan and H. Avanesyan, “Causes and conditions of occurrence of congestion situations in telecommunication networks”, *Computer Science and Information Technologies. Int. Conference*, September 25 – 29, Yerevan, Armenia, pp. 441--442, 2017.

[4] W.Stallings, *Foundations of Modern Networking*, by Pearson Education, Inc USA, 2016

[5] О. И. Шелухин и др., *Самоподобие и фракталы* , Физико-математическая литература, Москва, 2008.

Submitted 09.10.2017, accepted 06.12.2017.

Հեռահաղորդակցական ցանցերում գերբեռնվածքային իրավիճակները որպես հատուկ տիպի վթարային իրավիճակ և դրանց մշակման մեթոդները

Հ. Հարությունյան և Հ. Ավանեսյան

Ամփոփում

Գերբեռնվածքային իրավիճակների դեմ պայքարի խնդիրը հանդիսանում է հեռահաղորդակցության ցանցերի առավել բարդ և կարևոր խնդիրներից մեկը: .Այդ խնդրի ուսումնասիրմանը նվիրված են բազմաթիվ գիտական աշխատանքներ: Ներկայումս հետազոտությունները հիմնականում տարվում են հետևյալ ուղղություններով՝

- հեռահաղորդակցական կենտրոնների (հանգույցների) արտադրողականության բարձրացում
- կապուղիների թողունակության բարձրացում
- տվյալների փոխանցման արձանագրությունների կատարելագործում
- երթուղիների կառավարման արդյունավետ ալգորիթմների մշակում
- գերբեռնվածություն իրավիճակների կանխարգելման, հայտնաբերման, մշակման և վերացման նոր մեթոդների մշակում:

Հոդվածում հեռահաղորդակցական կենտրոններում գերբնակվածության իրավիճակները դիտարկվում են որպես հատուկ տիպի վթարային իրավիճակներ և դրա հետ կապված՝ գերբնակվածության դեմ պայքարի մեթոդները դիտարկվում են վթարային իրավիճակների դեմ գոյություն ունեցող մեթոդաբանության համատեքստում; Նման մոտեցումը ընդլայնում է գերբեռնվածության իրավիճակների ուսումնասիրության և մշակման հնարավորությունները:

Перегрузка как аварийная ситуация и методы ее обработки

Г. Арутюнян и А. Аванесыан

Аннотация

Борьба с перегрузками является одной из сложных проблем телекоммуникационных сетей. Решению этой проблемы посвящено много работ. В настоящее время исследования ведутся в направлениях увеличения производительности коммутационных центров, пропускной способности каналов связи, совершенствования протоколов передачи данных, создания эффективных алгоритмов управления трафиком, разработки методов предотвращения, обнаружения и устранения перегрузки с возможно минимальными потерями. Все эти исследования имеют цель поддерживать баланс между входящими и исходящими потоками коммутационного центра.

В данной работе перегрузка рассматривается как аварийное состояние в телекоммуникационных сетях. Общая стратегия управления аварийными ситуациями предполагает использование следующих процедур:

- предотвращения возникновения аварийной ситуации;
- контроля и обнаружения возникновения аварийной ситуации;
- обработки аварийной ситуации и восстановление нормального режима работы системы;
- устранения последствия аварийного состояния.

В соответствии с этой постановкой, методы борьбы с перегрузками в работе рассматриваются как методы борьбы со специальным видом аварийных ситуаций. Рассмотрение состояния перегрузки как особый вид аварийных ситуаций расширяет методологию исследования и борьбы с перегрузками, используя существующие методологии борьбы с аварийными ситуациями.