MODELING OF LTE EPS WITH SELF-SIMILAR TRAFFIC FOR NETWORK PERFORMANCE ANALYSIS

An important task in parametric synthesis of telecommunications network such as EPS net-work is to determine QoS parameters when traffic processed on EPS network nodes. In this paper, presented mathematical models for the calculation of QoS parameters such as delay and loss probability taking into account properties of self-similar traffic. This modeling based on model P/P/1 with Pareto distribution function for between packet intervals and processing time for case then intervals between packets and the duration of processing on servers are limited. Proposed model compared with model with log-normal and Weibull distribution of packet processing time and M/M/1 model.

Keywords: LTE, EPS, model, self-similar traffic, delay, drop probability, QoS.

Introduction

LTE stands for Long Term Evolution and it was started as a project in 2004 by telecommunication body known as the Third Generation Partnership Project (3GPP). LTE evolved from an earlier 3GPP system known as the Universal Mobile Telecommunication System (UMTS), which in turn evolved from the Global System for Mobile Communications (GSM). Long Term Evolution (LTE) describes the standardization work by the Third Generation Partnership Project (3GPP) to define a new high-speed radio access method for mobile communication systems. Rapid development of information technologies and their wide dissemination puts high requirements to telecommunication systems. We can satisfy these high requirements for telecommunication systems not only through the development of network management methods, but through the development of design methods.

The core network of the LTE-Advanced system is separated into many parts. Fig. 1 shows how each component in the LTE EPS network is connected to one another.

![Fig. 1. The EPS network elements](image)

The Mobility Management Entity (MME) is the control-plane node of the EPC. The Serving Gateway (S-GW) is the user-plane node connecting the EPC to the LTE RAN. The Packet Data Network Gateway (PDN Gateway, P-GW) connects the EPC to the internet. In addition, the EPC also contains other types of nodes such as Policy and Charging Rules Function (PCRF) responsible for quality-of-service (QoS) handling and charging, and the Home Subscriber Service (HSS) node, a database containing subscriber information. It should be noted that the nodes discussed above are logical nodes. In an actual physical implementation, several of them may very well be combined. For example, the MME, P-GW, and S-GW could very well be combined into a single physical node. The LTE radio-access network uses a flat architecture with a single type of node – the eNodeB, which response is responsible for all radio-related functions in one or several cells.

Essential to ensure the effective functioning of EPS (including LTE RAN) is efficiency of network planning. It was from design methods, the adequacy of
mathematical models used in this case, depend the properties and the viability of the future system.

In order to successfully compete to other existing and future wireless, cellular and wire-line services, the network designers need to fully consider the technical constraints that influence the whole design process of this kind of networks. The number of combinations of network elements and parameters that can be configured (e.g. antenna tilt, azimuth, base station location, power) constitutes the solution space of the design process. The size of this space determines the degree of complexity of finding appropriate solutions.

There are many different approaches for LTE network planning. Some papers are concerned about the radio part and its parameters’ optimization such as power allocation, radio resource scheduling, antenna down-tilt and BS positioning. Some papers are concerned about the radio part and its parameters’ optimization such as radio resource scheduling, antenna down-tilt and BS positioning [1, 2] or even traffic capacity planning approach for LTE radio networks [3]. In general, LTE network planning involves a myriad of components such as antenna height, antenna inclination angle, Base Station (BS) transmit power, BS capacity, BS position and transmission bandwidth.

In [1], Li et al. tackled two important components namely BS positioning and BS power allocation in LTE networks. The method used for locating the BS position and allocating the initial power is the service search method that is based on the traffic in the planning region; the desired BS position is calculated and taken into consideration if the traffic achieved in the coverage area of the biggest BS radius is less than the maximum load and more than the minimum load. If the coverage rate of the covered traffic doesn’t meet the requirement, then a smaller radius is chosen.

The LTE radio network capacity planning approach proposed in [3] is an iterative process that aims to find the optimal capacity planning solution taking into consideration specific requirements and parameters. There are two different types of parameters: basic engineering and radio parameters, and optimization parameters. The former includes different parameters such as transmission power, and system bandwidth; whereas the latter considers issues such as antenna down-tilt, distance between sites, etc. The unified traffic process module converts the complex various traffic requirements into uniform information that takes into consideration QoS requirement and the number of users for every traffic type.

Modern research effort traffic in telecommunication network [4, 5] show that its statistical characteristics are different from those adopted in the classical queuing theory. This leads to the fact that traditional methods of calculating the parameters of telecommunications systems and their probability-time characteristics, based on the Poisson model and the Erlang formula gives unduly optimistic results, leading to an underestimation of the load. Recent studies of the properties of traffic in modern networks have shown that the use of models of self-similar processes can more accurately describe the traffic transmitted in these networks.

Network performance degrades gradually with increasing self-similarity. The more self-similar the traffic, the longer the queue size. The queue length distribution of self-similar traffic decays more slowly than with Poisson sources. However, long-range dependence implies nothing about its short-term correlations which affect performance in small buffers. Additionally, aggregating streams of self-similar traffic typically intensifies the self-similarity (“burstiness”) rather than smoothing it, compounding the problem.

Because streams that transmitted through the network has properties of self-similarity, this exerts a great influence on the performance of the EPS. Particularly important role it plays in parametric synthesis of EPC network, providing the transmission of multimedia traffic and real-time traffic. Use for modeling of traffic in EPS models of self-similar processes in order parametric synthesis, highlights the need to address these particular problems:

- choice of mathematical models of traffic in different parts and levels of EPS network;
- development of methods for determining the parameters of the aggregate traffic, formed by combining traffic arriving at a processing node or transmitted together on a common links;
- determination of the calculated expressions that allow to relate the quality of service parameters with the parameters passed and served flows.

In this article, we consider the solution of such problems as determination of the calculated expressions that allow to relate the quality of service parameters for individual traffic in EPS network nodes, which served requests from LTE subscribers. These studies are an important part of the planning method for LTE and EPS networks, which will be considered in our future research.

Quality of Service Parameter Estimation for Self-Similar Traffic

In time of LTE RAN and EPC network capacity design is often necessary to determine the quality of service parameters for traffic generated by individual subscribers. Requests from users in this case are transmitted through the network and received for processing to the server. Requests incoming process can be considered as self-similar, or in some cases - as a Poissonian. Query processing time is a random variable and in most cases, the distribution is described as "heavy tail". In this case are not suitable the results
known from the classical teletraffic theory. However, for solving the LTE RAN and EPC network capacity design problems is necessary to find expressions which bind traffic parameters with probabilistic and temporal parameters of service requests.

In modeling of traffic in the network is often used such laws as the Pareto distribution, Weibull, Gamma distribution or log-normal. Among the most commonly used law is the Pareto distribution. This distribution is characterized by the parameter $\alpha$. If $1 < \alpha < 2$, then this distribution has a "heavy tail" and has infinite variance. In this case, the dispersion index $\zeta^2$ approaches infinity. However, in actual LTE EPC networks intervals between packets and the duration of processing on servers are limited. In this article we propose to use a limited distribution that allows not changing the shape of the "tail" (that adequately simulate the property of self-similarity) indicates the maximum interval between packets and duration of the processing request in the server. It provides limited value of dispersion index $\zeta^2$.

The coefficient of dispersion in the case of restrictions of a Pareto distribution can be defined as

$$\zeta^2 = \frac{(1-\alpha)^2}{\alpha} \left( \frac{L^\alpha - L^k}{(1-\alpha)^2 (L^\alpha - L^k)} \right) \cdot 2 \frac{L^\alpha - L^k}{L^\alpha - L^k},$$

where $L$ and $k$ – respectively, the maximum and minimum possible value of the length of time intervals.

If the packet arrival process is different from the Poisson and the distribution of service time - from the exponential (type system $G/G/m$), you can use the results of the theory of diffusion approximation [6]. The mean number of requests in the system can be determined [8] as

$$N = \rho \cdot \frac{1 + \zeta^2_{\text{in}}}{2(1 - \rho)}.$$  

Using (7), with the use the formula Little obtain an expression to determine the mean residence time packet in the system, described by model M/G/1

$$\tau = \frac{1}{m(1-\rho)} \cdot \frac{1 + \zeta^2_{\text{in}}}{\zeta_{\text{out}}^2}.$$  

where $\zeta_{\text{in}}$ – buffer size;

$\rho$ – system utilisation.

Another possibility is when the server receives the flow of requests that can be described by a Poisson process, and the processing time is not distributed exponentially. This case can be viewed as a queuing system $M/G/1$.

Accordingly, the formula Poliachek-Khinchin, the mean number of requests in the system can be determined [8] as

$$N = \rho + \rho^2 \cdot \frac{1 + \zeta^2_{\text{in}}}{2(1 - \rho)}.$$  

Using (7), with the use the formula Little obtain an expression to determine the mean residence time packet in the system, described by model M/G/1

$$\tau = \tau_\text{w} + \tau_\text{s}.$$  

Let us analyze proposed models in the article and compare the values of the parameters of quality of service traffic in the network nodes.

Analysis of the results of analytical modeling

In [7 – 10] for the calculation of packet loss probability in the $G/G/1/1$ queuing systems with known distributions laws for input traffic and servicing time, proposed to use the following approximate formula obtained based on the diffusion theory

$$P = \frac{1 - \rho}{1 - \rho \cdot \rho^2 \cdot \frac{1}{\zeta_{\text{in}}^2 + \zeta_{\text{out}}^2}},$$

where $\ell$ – buffer size;

$\rho$ – system utilisation.

As you can see from (5), the magnitude of losses in the system depends on the coefficient of variation, seen as the main characteristics of the flow and service process.

Let us consider some special cases. So for the system $M/M/1$, assuming that the interval between the packet arrival and packets servicing time is exponentially distributed, dispersion coefficients $\zeta_{\text{in}}^2 = 1$ and $\zeta_{\text{out}}^2 = 1$ and expression (5) degenerates to the known classical theory of teletraffic result

$$P = \frac{(1 - \rho)}{1 - \rho \cdot \rho^2}.$$  

In [7 – 10] for the calculation of packet loss probability in the $G/G/1/1$ queuing systems with known distributions laws for input traffic and servicing time, proposed to use the following approximate formula obtained based on the diffusion theory

$$P = \frac{1 - \rho}{1 - \rho \cdot \rho^2 \cdot \frac{1}{\zeta_{\text{in}}^2 + \zeta_{\text{out}}^2}},$$

where $\ell$ – buffer size;

$\rho$ – system utilisation.
At the beginning we investigate the packets delay in node for traffic which has interval between packets and processing time in the node described by Pareto distributions (model P/P/1). Results of research as a plot waiting time vs. node utilization for various values L, and for M/M/1 model shown in fig. 2, a.

Analysis of the results indicated that the traffic model described by P/P/1 has greater delay value than in the case of the M/M/1 traffic. In addition, when increasing the value of L, then increases waiting time in the node.

Next, we make a comparative analysis of packet delay in the network node in the case of traffic that describes different distributions of service time in the node network. Results of research as a plot waiting time vs. node utilization for such processing time distribution as the lognormal, Pareto, Weibull, as well as the model M/M/1 shown in fig. 2, b. Analysis of the results shows that the maximum delay is exposed traffic with log-normal distribution of the service time. Traffic Weibull distribution is less in case than the delay in the case of a Pareto distribution. The minimum delay for the case of traffic will be described by the model M/M/1.

Next, we make a comparative analysis of packet drop probability in the network node in the case of traffic that describes different distributions of service time in the node network.

Results of research as a plot packet drop probability vs. node utilization for such processing time distribution as the lognormal, Pareto, Weibull, as well as the model M/M/1 shown in fig. 3.

The experimental results are similar to the previous one case. So the maximum packet drop probability will be in case then traffic has lognormal distribution of processing time.

Traffic with Weibull distribution has a lower packet drop probability than in the case of the Pareto distribution.

The minimum packet drop probability in the case of traffic which described by M/M/1 model.

**Conclusion**

To maintain its competitive edge in the world of mobile networks in the future, 3GPP has initiated work on LTE. LTE is a packet optimized radio access technology with low latency and large bandwidths.

Essential to ensure the effective functioning of EPS (including LTE RAN) is efficiency of network planning.

Modern research effort traffic in telecommunication network show that its statistical characteristics are different from those adopted in the classical queuing theory.

Recent studies of the properties of traffic in modern networks have shown that the use of models of self-similar processes can more accurately describe the traffic transmitted in these networks.

Network performance degrades gradually with increasing self-similarity.

The more self-similar the traffic, the longer the queue size.
An important task in parametric synthesis of telecommunications network such as EPS network is to determine QoS parameters when traffic processed on EPS network nodes.

In this paper, presented mathematical models for the calculation of QoS parameters such as delay and loss probability taking into account properties of self-similar traffic.

Estimates of the probability of losses and delays for systems G/G/1 derived from the diffusion approximation can be considered adequate process models to EPS.

These studies are an important part of the planning method for LTE and EPS networks, which will be considered in our future research.

Reference